CHEMICAL RESEARCH, DEVELOPMENT & ENGINEERING CENTER

CRDEC-CR-88059

ASSESSING ECOSYSTEM IMPACTS FROM SIMULANT AND DECONTAMINANT USE

by H. W. Kerster
D. J. Schaeffer
K. A. Reinbold

U.S. ARMY CONSTRUCTION ENGINEERING RESEARCH LABORATORY
Champaign, IL 61820-1305

May 1988

U.S. ARMY ARMAMENT MUNITIONS CHEMICAL COMMAND

Aberdeen Proving Ground, Maryland 21010-5423

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

Distribution Statement

Approved for public release; distribution is unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CRDEC-CR-88059	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION USA-CERL 6b. OFFICE SYMBOL (If applicable) CECER-IMT	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) P.O. Box 4005 Champaign, IL 61820-1305	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION CRDEC 8b. OFFICE SYMBOL (If applicable) SMCCR-RST-E	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MIPR No. 84-1420
8c. ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground, MD 21010-5423	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Assessing Ecosystem Impacts from Simulant and Decontaminant Use

12. PERSONAL AUTHOR(S)

Kerster, H. W.; Schaeffer, D. J.; and Reinbold, K. A.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Contractor	FROM 85 Mar TO 85 Sep	1988 May	63

16. SUPPLEMENTARY NOTATION

COR: Dr. R. Wentsel, SMCCR-RST-E, (301) 671-3098

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
24	07	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Pollution; Risk assessment; Ecotoxicology; Simulants; Decontaminants; Contamination; Training; Environmental protection

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Chemicals for simulation and decontamination of chemical welfare agents may alter the ecosystems of training sites and their environs. Little information is available on the ecotoxicology of those chemicals. This report presents methods from the literature and proposes a newly devised approach to rank those chemicals for damage potential. Further data necessary to reduce ranking errors are identified, and chemical and ecological monitoring of training sites and their surroundings are recommended. The information in this report should be of special interest to U.S. Army personnel responsible for environmental protection of training areas. Keywords: Chemical warfare; Chemical agent Simulants; Environmental protection; Risk assessment; Contamination; Ecotoxicity. (CPS-K)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL SANDRA J. JOHNSON	22b TELEPHONE (Include Area Code) (301) 671-2914 22c. OFFICE SYMBOL SMCCR-SPS-T

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

PREFACE

The work described in this report was authorized under MIPR No. 84-1420. This work was started in March 1985 and completed in September 1985.

The use of trade names or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This document has been approved for release to the public.

ACKNOWLEDGMENTS

The contributions of the following persons are gratefully acknowledged: for technical suggestions, Dr. Edwin Herricks, University of Illinois at Champaign-Urbana, Dr. Edward Novak, Team Leader of the Environmental Resources Team, USA-CERL, and Dr. Randall Wentsel, Project Monitor, CRDEC; for administrative support, Dr. Ravinder K. Jain, Chief, Environmental Division, USA-CERL. Dr. Kurt Enslein, Health Designs Inc., is thanked for permission to reproduce Appendix B.

COL Norman C. Hintz is Commander and Director of USA-CERL, and Dr. L. R. Shaffer is Technical Director. BG Peter D. Hidalgo is Commander of CRDEC.

Blank

CONTENTS

TABLES

FIGURES

Blank

ASSESSING ECOSYSTEM IMPACTS FROM SIMULANT AND DECONTAMINANT USE

1. INTRODUCTION

Simulants are substances for which the properties resemble those of other compounds enough to be used to mimic or study specific attributes of those compounds. Simulants are important in Army chemical defense programs because they can replace more hazardous chemical agents in training personnel and developing chemical defense materiel. The ideal chemical agent simulant would mimic all important chemical and physical properties of an agent (e.g., vapor pressure, solubility, reactivity) without having its undesirable environmental attributes. Decontaminants are substances that neutralize chemical agents on personnel and equipment.

To support environmental documentation for testing and training with these agents, methods are needed to project effects based on existing data if possible[1-3]. If existing data do not suffice, the methods should identify the additional data needed. This report presents models that use varying amounts of data for ranking ecotoxicological effects of chemicals.

The generic problem is: given particular chemicals, use scenarios, and natural communities, how can you project the ecosystem impacts of chemical use? Also available is common laboratory data (physical and chemical properties and laboratory animal acute toxicity), but no data on behavior of the substances in natural media or effects on members of the natural communities.

The problem is certainly a common one and will continue to be, for it must be solved by those who prepare and review environmental impact reports. This report surveys, largely from existing data, solutions in the literature and proposes a new one devised especially for projecting the effects of training use of simulants and decontaminants.

Models should be inexpensive to use, while assuring low rates of misranking as either false positives or false negatives. They should be individually validatable against already known compounds and their effects. A good model would require no more expertise to operate than is already present in personnel apt to use the model. If possible, the model should be implementable in an expert system computer program that minimizes operator effort and time, and that reduces the opportunity for operator error.

1.1 Approach and Scope.

The conceptual basis for the approach is identification of toxicological effects that have a high probability of causing appreciable harm to at least one major ecological system component. If toxic effects are unlikely, dispersion scenarios are not considered. When toxic effects are likely, the Army's already developed Preliminary Pollutant Limit Value (PPLV)[4] dispersion models are used to identify those ecosystem compartments having hazardous loads. We make no attempt to identify specific target species, adjust for local features, nor specifically quantify the effects on the target species or the entire ecological system. If, for example, there is evidence that a specific compound is acutely toxic to laboratory mammals, we assume that it is also acutely toxic to mammals in the ecological system of concern. We are not, in general, able to project across broader taxonomic boundaries from standard

toxicity data. The ecological system endpoints considered are acute and chronic toxic effects produced in mammals, fish, and plants.

The limited list of ecological endpoints is a small subset of possible effects. We justify the limitation as follows. The interdependence of components in ecosystems is known to be great, but pathways and interaction details are poorly understood. The effects of small amounts of xenotoxic chemicals (as opposed to bulk nutrients) on pathways and interactions are especially vague. This ignorance arises from the usual study of intact natural ecosystems not challenged by xenobiotics. The difficulty is well illustrated by acid rain, where new effects and pathways are continually being discovered, despite the native and bulk nature of the damaging chemicals.

Our approach focuses on conspicuous structural features having analogs of economic and human health interest. Data from the analogs are relatively common. These data can be projected to ecosystem analogs fairly safely, so that findings for the subset are robust, especially in light of the great uncertainty in further projection.

2. RISK RANK MATRIX METHOD

Very little toxicological data are available for the products considered here. As we extrapolate or estimate more and more properties from these limited data, errors propagate rapidly, increasing the chance of misclassifying a compound's expected ecological system effect. Similar reasoning has led to the equally crude ranking system for hazardous waste sites (the MITRE Corporation's model[5]). Our evaluation of the problem suggests that effects on an ecological system depend on three primary factors: the toxicity of the compound(s), the exposure levels used, and the effective biological contact area (EBCA). In our scheme, each of these factors can take on values of "low," "medium," or "high," or numerically, 1, 2, or 3. The product of these three scores provides a relative measure of ecological system hazard. The possible scores and corresponding ranks are:

Score	1, 2, 3, 4, 6, 8, 9, 12, 18, 27.
Rank	1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

The ecological system hazard ratings developed in this way are given in Table 1 as a "risk rank assignment matrix". As an example of how to use this table, a score of 6 is achieved if one factor is "low," one is "medium," and one is "high." We thus consider the ecological effect to be equivalent when a highly toxic compound (3) is used at low exposures (1) over a moderate EBCA (2), and when a low-toxicity compound (1) is used at a moderate level (2) over a large EBCA (3). Features such as decomposition of products, media transfer and accumulation, and bioaccumulation are modifying factors that can be included if they are known or can be reasonably estimated. Ranks for entry to the matrix can be refined using the battery and decision tree models described below.

Table 1
Risk Rank Assignment Matrix

Toxicity	Low Exposure			Medium Exposure			High Exposure		
	Low EBCA	Med EBCA	High EBCA	Low EBCA	Med EBCA	High EBCA	Low EBCA	Med EBCA	High EBCA
Low	1	2	3	2	4	5	3	5	7
Medium	2	4	5	4	6	8	5	8	9
High	3	5	7	5	8	9	7	9	10

Acute toxicity ranks in human equivalents proposed by the National Research Council[6] (NRC) are:

Low > 500 mg/kg
Medium > 50 - 500 mg/kg
High < 50 mg/kg

Based on training scenario information and ecological system considerations (e.g., increases in numbers of pathways and possible target species as area increases), we propose EBCA designations as:

Low < 1 acre
Medium > 1 - 10 acres
High > 10 acres

Designation of "exposure" levels is difficult since exposure is the product of the concentration and quantity used scaled by the area. Expressing the toxicity in human equivalents (70 kg, also approximate for deer), we propose the following exposure ranks per square meter (m^2):

Low, 70 (LD50 or LC50)/100 to 70 (LD50 or LC50)/10 mg/m^2
Medium, 70 (LD50 or LC50)/10 to 70 (LD50 or LC50) mg/m^2
High, at least 70 (LD50 or LC50)

2.1 Toxicological Assumptions.

The theoretical basis for this approach lies in several assumptions that are commonly accepted by toxicologists. First, most complex toxicological responses arise through a series of identifiable (if unknown) steps. Some of these steps are highly predictive of the overall toxicological response. As a simple example, the toxic response resulting from chronic exposure to lead is loss of neurological function. Such loss can be predicted on the basis of lead levels in blood. Second, a response in one species often indicates that similar responses will be found in other species. For example, rodents are

used to predict the toxicity and efficacy of drugs before human trials are conducted. Third, the ability to predict toxic responses in a species of interest improves as the number of species exhibiting toxic responses to a given chemical increases. For example, a chemical which is a carcinogen in mice, rats, rabbits, and chimpanzees is more likely to be a carcinogen in man than is a chemical which is carcinogenic only in male rats. Fourth, no prediction is perfect since unusual interspecies differences occur. The extreme teratogenicity of thalidomide in humans, but not in other mammals, is an example. Fifth, the author's position is that, in the absence of experimental data, the toxicity of mixtures is predicted best by dose additivity.

These toxicological considerations are consistent with the two-level weight-of-evidence approach taken here. The lower level consists of individual tests and models, each of which is a prediction of ecological effect. At the higher level, the independent results taken together predict higher level responses. As a common example of this approach, consider the screening of chemicals for carcinogenicity. Because compounds which are carcinogens are also often mutagens, carcinogenicity is usually predicted by mutagenicity. In our context, "carcinogenicity" is an ecological system effect of interest and "mutagenicity" is a prediction of that response made by a particular model. Thus, the lowest level battery for identifying carcinogens consists of mutagenicity tests in several bacterial strains. In some tests, the compound is mutagenic; in others it is not. Comparably, the lowest level ecological "models" include solubility in water, vapor pressure, and prediction of response. Although generally not all of these responses agree, decision criteria can be developed for determining the overall prediction of the battery. If this evaluation suggests that the compound is mutagenic (in our analogy), mammalian cell and organ tests in the next battery are performed. Our comparable environmental concerns include transport and fate models. At the highest level, compounds are tested through chronic lifetime exposures in at least two species. Environmental models at this level might include population models.

2.2 Chronic Toxicity.

Chronic effects are the most important ecological endpoints since they are the most difficult to recognize. Commonly, the chronic exposure is expressed as a "chronicity factor" which is calculated using acute toxicity data. Quoting Stevens and Gallo[7]:

> The chronicity factor is the ratio of the one-dose LD50 (mg/kg) divided by the 90-dose LD50 (mg/kg/day) for any particular compound. A comparison of day 1 to day 90 LD50s can tell the toxicologist a great deal about the in vivo handling of the compound. If the 90-day LD50 is far below that of the single dose LD50, it is probable that the compound is slowly metabolized and accumulates in the body. Recording of organ weights (including brain weight), food consumption, body weight, behavior and water consumption can go a long way in defining target organ toxicity.

Chronic exposure at low dose to a material toxic at high acute dose is often harmless since low doses may not overwhelm detoxification mechanisms. Chronic effects are often mitigated by degradation. We suggest that a possible index of chronic effect is some function which is related to both persistence and exposure.

Extensively adapting Petrocelli[8], chronic toxicity is an important tool for understanding and evaluating the potential hazard of chemicals to organisms exposed for a lifetime. Data from chronic tests can be used to estimate the effect and no-effect concentrations of a chemical to which organisms are exposed continuously during an entire reproductive cycle. A chronic toxicity test can indicate the concentrations of a chemical that will interfere with normal growth, development, and attainment of reproductive potential of an organism. Generally, concentrations that produce chronic effects are lower than those that produce more readily observable acute effects such as mortality. Therefore, chronic toxicity tests can provide a more sensitive measure of chemical toxicity than can acute toxicity tests.

Chronicity testing is necessary since it is not usually possible to infer chronic effects from acute exposure data. Kenaga[9] extensively compiled data on acute and chronic toxicity of chemicals to fish and aquatic invertebrates. His study of the acute/chronic ratios (ACRs) for 84 chemicals revealed no general rules useful in the chemical agent context.

A chronic toxicity test is designed to expose all lifestages of the test animal--viable gametes, newly fertilized ova, early stages of developing embryos, or young--to a range of chemical concentrations estimated (from acute toxicity test exposures) to bracket the threshold for significant deleterious effects (Table 2). It is especially important to test stages in the reproductive process (Table 3). If, for example, a chemical chronically interferes with an animal's sense of smell, the deficit may only be important during courtship, selection of egg-laying sites, or to a plant reliably pollinated only by the animal with intact senses. If appropriate test concentrations have been selected, the populations exposed to the higher concentrations in this test will be adversely affected, as judged by standard criteria, while those exposed to the lower concentrations will not be adversely affected compared with unexposed populations (controls).

Table 2
Some of the Sublethal Effects of Pollutants on Life Stages of Various Animals (After Sheehan[10] and Waldichuk[11])

Life Stage	Vital Life Process	Critical Effects of Pollutants
Egg	Meiotic division of cells; fertilization; cleavage mitoses of fertilized egg; hatching; respiration.	Gene damage; chromosome abnormalities; damage to egg's membrane; direct toxicity to embryo from pollutant; impaired respiration; reduced hatch.
Larva	Metamorphosis; morphological development; feeding; growth; avoidance of predators, parasites, and disease.	Toxicity from bioaccumulated poisons in yolk sac during early feeding; biochemical changes; physiological damage; deformities; behavioral alterations.
Juvenile	Feeding; growth; development of immune systems, endocrine glands; avoidance of predators, parasites, and disease.	Direct toxicity; reduced feeding and growth; altered predator-prey relations; impaired chemoreception; reduced resistance to parasites and disease.
Adult	Feeding; growth; sexual maturation.	Direct toxicity; adverse alteration of environmental conditions, e.g., dissolved oxygen; physiological and biochemical changes; behavioral alterations.

Table 3
Some Effects of Pollutants on Reproduction (After Sheehan[10])

Vital Process	Critical Effects of Pollutants
Development of gametes	Incomplete or abnormal development of ova or spermatozoa; gene damage.
Fertilization	Interference with homing of spermatozoa to the ova; impairment of the ability of spermatozoa to enter the micropyle and fertilize the ova successfully.
Embryo development	Cytological and cytogenetic abnormalities including chromosome bridging, breakage and translocation; interference with hardening of the egg; interference with gas and water exchanges; cessation of development.
Hatching	Failure to hatch; high mortality of newly hatched larvae; teratogenic abnormalities.
Sexual maturation	Histopathological effects on gonads; changes in production and metabolism of gonadotropins.
Courting and mating	Destruction of spawning and mating grounds; inappropriate courting or mating behavior leading to reduced mating success.

2.3 Risk Assessment Definitions.

Risk assessment for humans has been the subject of a study by a Committee of the NRC[6]. Risk assessment in the present context is the use of factual data to define the effects on the health of ecosystems produced by simulants, decontaminants, carriers, and use formulations ("products") at the concentrations projected to be in the environment. For the purposes of this study, "ecological system" means the nonhuman organisms, the relationships among them, and the relationships between the organisms and nonliving objects and associated space. Risk assessment for ecological systems then consists of the elements below.

Hazard identification: determination of whether a product is likely to contain contaminants that may be causally linked to ecological system health.

Exposure assessment: determination of the extent of ecological system exposure to the product, taking actual use conditions into account.

Risk characterization: description of the nature and magnitude of ecological system risk (including attendant uncertainty) that could result from use of the product.

Risk management: use of risk assessment data in decision-making. In the present context, this use will involve development of the criteria for ranking a given product based on the risk characterization. The key to risk management is decision-making. It cannot be emphasized too strongly that there is always an element of uncertainty in such decision-making, and therefore, it should be done by a group of the most qualified experts available.

3. THE THREE-BATTERY PROGRESSION

We propose a three-battery-of-tests approach to project the effects of simulants and associated compounds on the environment. The terms "test" and "model" are used interchangeably. Battery 1 consists of literature reviews and simple computational models[12]. Battery 2 consists of simple, single-species, short-term bioassay screening tests and chemical analyses. In contrast to Battery 2, Battery 3 involves ecological system studies that are generally multispecies, expensive, and lengthy. Batteries 1 and 2 predict ecological system effects that can be explored through field studies in the third battery. This approach minimizes expense because harmless chemicals are eliminated from consideration early in the procedure. The statistics necessary for estimating the correlation between two or more batteries have been developed[13]. Error rates for a particular substance out of several substances tested, and error rates for the set of chemicals tested, can be estimated as shown in Appendix A. This approach allows the set of tests within a battery to be adjusted to produce a desired error rate.

3.1 Overview of Batteries.

As shown in Table 4, classical toxicology is generally concerned with human health-related criteria. These criteria can be classified as: (1) general toxicity (acute, subacute, subchronic, chronic); (2) carcinogenicity and reproductive toxicity; (3) other (acute skin and eye irritation, allergic sensitization, neurotoxicity, hepatotoxicity); and (4) human studies (observations in humans, health surveys/census statistics, epidemiological studies). The first two categories include the toxic responses of greatest public concern and are the first considerations for prioritizing risk management efforts. While a much higher level of demonstration of "safety" is required for genotoxic effects than for nongenotoxic ones, very much less reliable data are available for the genotoxic endpoints, and these data cost substantially more to obtain[14,15].

Table 4
Endpoints of Toxic Effects

Toxic Response Category	Data Availability, % of NRC[15] Sample
(1) Toxicity, General (Hierarchical Order):	
Acute toxicity	80
Subacute toxicity (28-day)	1
Subchronic toxicity (90-day)	1
Chronic toxicity	1
(2) Carcinogenicity/Reproductive Toxicity (Semihierarchical Order):	
Mutagenicity	<10
Teratogenicity	
Reproductive toxicity	<5
Carcinogenicity	5 - <10
(3) Toxicity, Other (Random Order):	
Acute skin and eye irritation	5
Allergic sensitization	
Neurotoxicity	
Hepatotoxicity	
(4) Human Studies (Hierarchical Order):	<5
Observations in humans	
Health surveys/census statistics	
Epidemiological studies	
(5) Ecological System Toxicity Effects (Random Order):	
Abundance and biomass	
Reduction in population size and extinction--	
Loss of species with unique functions	
Species richness	
Community composition and species dominance--	
Species lists	
Indicator species	
Biological indices	
Dominance patterns	
Species diversity and similarity	
Spatial structure	
Stability--	
Inertia	
Elasticity	
Amplitude	
Hysteresis and malleability	
Persistence	
Succession and recovery	

In contrast to health-related criteria, ecological system effects[10] generally are not included in classical toxicology. Thus, although data are severely limited for common laboratory species, there is generally no toxicological data for species of ecological system interest. The outstanding characteristic of ecological system toxicological endpoints is the need to evaluate subtle, complex effects concurrently in multiple species that are not normally studied by toxicologists. The responses of interest, such as changes in diversity, are difficult to define clearly and unambiguously. Generally, these responses have poor precision (high measurement variance) and unknown accuracy. Thus, while many health-related responses (1-4 in Table 4) are components of ecological system effects (5 in Table 4), they are seldom endpoints per se.

System complexity and data scarcity thus limit the toxicologist's ability to predict the magnitude and significance of toxic effects of chemicals on ecological systems. Although some models of ecological system effects have been proposed, they have substantial data requirements and introduce significant problems in interpretation. To avoid the problems associated with inherently complex models, we consider three hierarchical batteries of simple models to predict effects of toxic chemicals on ecological systems. We anticipate that most decisions regarding the acceptability of simulants, carriers, decontaminants, and use formulations will be made inexpensively with the first battery.

Battery 1 consists of models to estimate solubility in water, octanol-water partition coefficient, acute toxicity, soil adsorption, vapor pressure, volatilization from water, volatilization from soil, diffusion coefficients in air and water, rate of hydrolysis, rate of aqueous photolysis, and allometric relationships. Some of these models are stand-alone. For example, if the primary concern is acute toxicity to rabbits, and experimental data exist, the "model" presented as Table 11 can be used for decision-making. Other models are generally used hierarchically. Suppose we need to estimate the acute toxicity of a benzene derivative to rabbits knowing only the compound's molecular structure and solubility in water. We might first estimate the octanol-water partition coefficient and then use this value to estimate the acute toxicity to mice. Allometric relationships would then be used to convert the estimate of acute toxicity in mice to an estimate of toxicity in rabbits.

Battery 2 includes structure-activity relationship (SAR) models, methods to estimate carcinogenicity and mutagenicity in mammals, interspecies conversions, estimates of phytotoxicity, bioconcentration factor estimates for aquatic organisms, bioconcentration estimates for other species, and methods for estimating the total weight of active compounds in complex mixtures. For example, this battery considers a model to convert acute toxicity data in rodents to acute toxicity in bluegills.

Battery 3 includes multispecies and ecosystem models such as microcosms, connectedness, population structure and dynamics, diversity, measures of community metabolism, nutrient flux, taxonomic composition, and chemical disposition. Several of these features are listed in Tables 5 and 6.

Table 5
Ecosystem Attributes (After Odum[16])

	Community Energetics
1.	Gross production/community respiration (P/R ratio)
2.	Gross production/standing crop biomass (P/B ratio)
3.	Biomass supported/unit energy flow (B/E ratio)
4.	Net community production (yield)
5.	Food chains
	Community Structure
6.	Total organic matter
7.	Inorganic nutrients
8.	Species diversity--variety component
9.	Species diversity--equitability component
10.	Biochemical diversity
11.	Stratification and spatial heterogeneity (pattern diversity)
	Life History
12.	Niche specialization
13.	Size of organism
14.	Life cycles
	Nutrient Cycling
15.	Mineral cycles
16.	Nutrient exchange rate between organisms and environment
17.	Role of detritus in nutrient regeneration
	Selection Pressure
18.	Growth form
19.	Production
	Overall Homeostasis
20.	Internal symbiosis
21.	Nutrient conservation
22.	Stability (resistance to external perturbations)
23.	Entropy
24.	Information

Table 6
Measures of Ecosystem Dynamics and Examples of Their Application (After Holling[17], Westman[18], and Sheehan[10])

Characteristic	Definition	Example: Ecosystem Subjected to Oil Spill
Inertia (ecological buffering capacity)	Resistance to change.	Amount of oil that must accumulate over a given area in a given time period to cause a given level of ecosystem damage (such as local extinction of species X and Y).
Elasticity	Rapidity of restoration of a stable state following disturbance.	Time required to recover initial structure or function following ecosystem damage (e.g., restoration of populations X and Y).
Amplitude	Zone from which the system will return to its original state.	Maximum amount of oil that can accumulate in an area such that damage sustained can be fully repaired (e.g., restoration of populations X and Y).
Resilience	Zone from which the system will return to a stable configuration, perhaps different from the original.	Maximum pollutant level after which populations will stabilize at some level.
Hysteresis	Degree to which path of restoration is not an exact reversal of the path of degradation.	Degree to which pattern of secondary succession is not an exact reversal of the pattern of retrogression experienced following impact. (e.g., Were the last species to disappear the first ones to return?)
Malleability	Degree to which stable-state ecosystem established after disturbance differs from the original steady state.	Degree to which new climax resembles the initial climax state. (e.g., How closely do the species composition and equitability of the new climax state resemble the old?)

3.2 Battery 1.

3.2.1 Water Solubility, Hydrolysis, Octanol-Water Partition Coefficient.

Compounds that are freely soluble in water or that undergo substantial and relatively rapid hydrolysis generally do not bioaccumulate. Primary biological effects of such compounds are typically acute toxic effects such as cholinesterase inhibition.

3.2.2 Models for Acute Toxicity Using Octanol-Water Partition Coefficient.

For many chemicals, various toxic endpoints are well correlated with solubility through the octanol-water partition coefficient (log P). In the general case, the models postulate that molecular fragments (A, B, C . . . N) contribute an additive biological activity (BA):

$$BA = aA + bB + cC + \ldots + nN. \quad (1)$$

The "best" values of the numerical coefficients (a, b, c. . . n) are obtained from known compounds using weighted statistical methods. These relationships are relatively inexpensive to use once an appropriate model (Appendix B) has been developed since no chemicals are required for testing.

Nongenotoxic responses induced by chemicals arise through the two broad mechanisms of (1) "narcotic" or (2) "specific" toxicity. As proposed by Ferguson[19] in 1939, narcotic toxicity largely results from a reversible decrease in physiological function. Thus, he proposed that when a physiological effect is reversible, when an equilibrium exists between the organism and the external phase, and when the physiological effect is a function of the external concentrations, the toxic effect is primarily physical (rather than chemical) in character. Mullins[20] divided these effects into narcosis of cell division and narcosis of the central nervous system (general anesthesia). For example, many classes of organic compounds, such as organic solvents and chlorinated hydrocarbons, appear to act through narcotic toxicity mechanisms. Generally, the toxicity of these compounds is proportional to their water solubility as measured by P, the octanol-water partition coefficient.[21]

$$\log (1/C) = -a(\log P)^2 + b \log P + k. \quad (2)$$

In (2), C is the concentration producing a standard biological response, such as LD50. For many nonionizable organic compounds, the partition coefficient also gives a good indication of bioaccumulation potential[22]. It has been found for aquatic organisms that when $P < 1000$, the estimated bioconcentration factor (BCF, wet weight basis) is below 100.

At small values of log P, the equation may appear linear. However, as the range of log P is extended, the relationship between biological activity and log P reverses and toxicity actually decreases with increasing log P. Since log P and water solubility are inversely but linearly related, the expected physical toxicity for chemicals with a large log P could be above the expected water solubility and no effects would be measured. This "parabolic effect" can be described more accurately as a bilinear effect resulting from competing membrane transport phenomena[23,24]. The bilinear model[21] (3) is more consistent with the understanding of membrane transport, which suggests that

water-soluble-chemical transport may be controlled by the membrane whereas lipid-soluble-chemical transport may be controlled by the diffusion layer.

$$\log (1/C) = a \log P - b \log (BP + 1) + k. \qquad (3)$$

3.3 Battery 2.

Gillett[25] developed a protocol for deciding which chemicals should be subjected to what tests to minimize both testing costs and ecosystem risk. His scheme depends on inferences from the types of limited data we find on environmental data sheets and structure-activity relationships that are derived from such data. Gillett's scheme emphasizes ecotoxicological inferences that can be drawn from simple combinations of chemical properties. For example, he argues that chemicals for which the logarithm of the octanol-water partition coefficient plus 1.631 times the half-life (in days) exceeds 5.37 might be especially dangerous, since they are soluble enough to reach biota, insoluble enough to accumulate in lipid stores once in biota, and of long enough half-life so that biota are not protected from exposure by degradation of the chemical.

Like our risk rank matrix method, Gillett's screening protocol assigns chemicals to one of three ranks (heavy, low, or no concern). His method is the best of the models we have found for use with existing simulant and decontaminant data. He does not, however, rank products.

Aquatic systems have been the subject of most modeling efforts, whether the targets of concern have been human or nonhuman. The TOXSCRN1 model[26] is a lake management aid. From the water-sediment partition coefficient, decay rates in those media, diffusion coefficients and physical features of a lake, the model returns chemical concentrations and related variables more directly related to impact on biota than the entry variables. A microcomputer program is available that eliminates manual calculation.

Similar to TOXSCRN1 is a USEPA model[27] that also yields compartment concentrations. In addition, it allows as input data on biological, as well as physical, transformation rates.

A human fate and effects computer-based modeling system from USEPA, the Graphical Exposure Modeling System (GEMS), is worth noting here. One component of GEMS, the Exposure Analysis Modeling System (EXAMS), permits estimation of partition coefficients from standard chemical data[28].

3.3.1 Models Incorporating Electronic and Molecular Parameters.

Simple models based only on solubility fail where "specific" toxicity is exhibited by narcotic toxicants (at high exposures) and by certain classes of compounds (such as organophosphate and carbamate pesticides) at virtually all exposures. Where a homologous series of compounds is available, the Hansch model (Eq. 4 and Veith[21]) is generally very powerful, and the constants a, b, c, d, and e, are readily obtained using multiple regression. In addition to the Hansch-Fujiti "pi" value estimate of the partition coefficient, this model includes terms for substituent electronic effects on the electron density at a remote center (Hammett "sigma"), steric effect (Taft "E" value), and the molar refractivity (MR) of substituents. Usually a parabolic term is included to

adjust for the fact that the relationship between log P and log (1/C) is nonlinear.

$$\log (1/C) = a + b(pi) + c(sigma) + d(E) + e(MR). \quad (4)$$

3.3.2 Models for Genotoxic Effects.

Equation (4) can also be used to estimate genotoxic effects. For example, for seventeen 1-(X-phenyl)-3,3-dialkyltriazines, the molar concentration necessary to give 30 revertants per 10^8 Salmonella typhimurium TA92 in the Ames test[29-31] is:

$$\log (1/C) = 1.09 \log P - 1.63(sigma+) + 5.58, \quad (5)$$

where "sigma+" is the "through resonance" electronic parameter. Purcell[31] notes that when (5) is compared with the antitumor activity[32] of the triazines against L1210 leukemia and general toxicity (LD50) in mice[29], it appears that mutagenicity is more sensitive than antitumor activity to the electronic effects of substituents, and that both gross toxicity and antitumor activity in mice show the same dependence on the electronic effects of substituents.

3.3.3 Models for Lethality, Carcinogenicity, Mutagenicity, Teratogenicity.

The simple models described above are inadequate if compounds with diverse structures and functional groups are considered. More complex models consider hundreds of structural features and complex statistical methods are needed to identify the several dozen significant ones. There are two aspects of developing and then applying these complex relationships: classification and quantification. The classification problem is that of identifying which compounds in a group of compounds are, for example, carcinogens and which are noncarcinogens in a standard rodent bioassay. The multivariate statistical methods of factor and principal component analysis, pattern recognition, and discriminant analysis have been used extensively to develop classification criteria for various endpoints. The quantification problem is often not separated from classification. For example, quantitative estimates of potency for individual compounds have been made using stepwise multiple regression. Discriminant analysis has been used to classify and quantify rat oral LD50[33] mutagenicity based on the Ames test[34], carcinogenicity[35], and teratogenicity[36]. If data on the metabolic pathways and metabolites are incorporated into structure-activity relationships, the models can be used to predict metabolic pathways and metabolites of new compounds.[37]

Most of these models have been published in the open literature. However, applying them to a new compound requires a sophisticated understanding of both chemistry and statistics, so the certifying organization should be responsible for having these computations performed at the time of submission. For example, a predictive model for carcinogenicity uses over 80 structural features[35]. For the 343 compounds studied[35], the model has a false negative rate (i.e., those definite carcinogens which the discriminate equation evaluates as indefinite carcinogens) of 3.6 to 4.9 percent. Conversely, the false positive rate (i.e., compounds identified by animal testing as non- or indefinite carcinogens) is 10.8 to 11.7 percent. Another way of looking at the power of the discriminant function is to note that between 87 and 91 percent of the definite carcinogens are correctly classified by the equation

and that between 77.5 and 80 percent of the non- or indefinite carcinogens are so classified.

3.4 Battery 3.

Models in this battery are those for which the data requirements are most costly in both elasped time and money. These models require as input Battery 2 output and also need ecosystem state variables (population or trophic level sizes) and/or ecosystem process variables. Fortunately, recent years have seen the development of models with light demands for ecosystem data and some ability to fill in the gaps and deliver ecosystem impact estimates for products.

A recent survey of ecosystem impact case studies[38] identified several ecosystem variables that can be estimated fairly inexpensively and that have been found to cause or accompany ecosystem damage from pollutants. The size of the nutrient pool tends to increase in aquatic and decrease in terrestrial ecosystems. Those changes are determined easily from chemical and volume measurements on streams that drain training areas. Primary productivity, not as easily measured, displays the same responses as the nutrient pool. Species diversity tends to decrease in pollutant-impacted communities, as does size variability among community members. The reduction in size variability is often due to loss of the larger members (large fish or trees). Pollutant stressed ecosystems are often invaded by species from earlier successional stages, i.e., the system appears to regress successionally. Regression is monitored easily by scoring the relative importance of organisms common in the preceding successional stage.

Hakanson[39] has attempted a conceptual framework concerning aquatic contamination and ecological risk. He formalizes the relationship between exposure, recipient sensitivity, and potential effect as:

$$E = f(D, T, W_i) + R, \qquad (6)$$

where:

E = a parameter expressing ecological effect
D = an exposure parameter
T = a factor expressing toxicity
W_i for i = 1, 2, 3, 4 . . . n = factors expressing recipient sensitivity
R = a residual term (the unaccountable remainder).

In this way, the potential ecological effect (E) is a function of the exposure (D, which could be given as a concentration or a load of a substance or a wastewater), the toxicity of the contaminant (T), the sensitivity of the recipient to this given substance or effluent water (W_i), and a residual term R, which expresses the fact that it is practically impossible to establish a 100 percent explanatory model in ecological contexts from a limited number of variables. He stresses that the crucial point with this approach, like ours, is to quantitatively express normative E-values from a limited number of readily available, inexpensive and representative integrating variables.

The models considered so far solely identify the type of effect, and perhaps its relative magnitude, on a target species. For example, a compound

may be identified as being toxic to plants but not to mammals or insects. We now make explicit the fact that the target organism in these models is not necessarily a single species but may, and usually will be, a group of organisms with similar characteristics. For example, if "deer" is a target organism, in the broader ecological sense "deer" may be treated as surrogates for large mammalian herbivores, such as sheep and cattle, using the same food sources. As another example, if our concern is substances causing deoxygenation of a small stream, the appropriate target "organism" is all freshwater invertebrates demanding high (> 5 mg/L) dissolved oxygen levels, not just, e.g., the daphnia used in the bioassay.

Aggregration of species into trophic levels must be done cautiously, however. The SWACOM lake simulation model[40] found that estimates of risk that included population-specific toxicities were two to three times the risks estimated from trophic toxicities. Thus aggregation into guilds or trophic levels may obscure deleterious effects on particular members of the aggregrate that are of interest to the Army.

Leaving aside the issue of aggregration, which must be dealt with in the context of particular ecosystems, the task now is to translate qualitative or quantitative estimates of hazard to selected target organisms into significant effects in the ecological system. To do so, we must move away from the one-link direct-effects models--generally considered in bioassay testing and by population-theoretic ecologists--toward the systems analyses and network interactions appropriate to actual ecological situations.

Two systems of models[41-46] based on matrix algebra approaches to system connectivity are presented here. These models, and their underlying mathematical theories, are very complex, and a detailed description is beyond the scope of this report. These models can be used appropriately only by an ecologist because they require intimate ecological knowledge and sophisticated mathematical understanding. The first example considers (in summarized form) a bioenergetics model for an oyster reef ecosystem[46]. The second example is a simplified model of an insecticide-dosed terrestrial ecosystem[44].

Figure 1 shows Patten's model[46] of an oyster reef ecosystem. The filter feeding compartment (1) consists of the American oyster (Crassostrea virginica) and a mussel (Brachidontes exustus). Deposited detritus (2) comprises the feces and pseudofeces produced by the feeding and excreting mechanisms of the shellfish. Microbiota (3) consists of bacteria, yeasts, and fungi associated with detritus. The meiofauna (4) are defined as benthic animals which pass through a 1-mm sieve, but are retained by a 0.063-mm sieve. Deposit feeders (5) consist of macrofauna that feed in the sediments. Predators (6) are the animals directly benefitting from this mode of interaction (predator-prey) with other forms. The single system input consists of phytoplankton and suspended detrital particles acquired by respiration, mortality, and resuspension. The energy flow processes that couple these compartments internally, and also those which provide input and output linkage to the environment, are detailed in Dame and Patten[47]. A qualitative description of the path analysis for this system follows.

3.4.1 Path Analysis.

Let $A = (a_{ij})$ be the adjacency matrix (Table 7) that represents the connectivity of the system in the figure. This matrix has $a_{ij} = 1$ if a direct causual interaction (energy flow) exists from compartment j (column) to compartment i (row), and $a_{ij} = 0$ otherwise. Each $a_{ij} = 1$ denotes a directed path of length 1 (j ->1) in the model. If A is multiplied by itself, the product A^2 indicates the number of indirect paths of length 2 (j->k->i) from compartment j to compartment i (Table 8). In general, a product matrix A^L represents the number of length L paths from j to i. For L = 0, A^L equals the identity matrix, I. In a system, the total number of incoming paths of length L to compartment i is the sum of the ith row of A^L, the total number of outgoing paths of length L from compartment j is the sum of the jth column of A^L, and the total number of paths is the sum of the outgoing and incoming paths taken over all L lengths. Patten's most striking conclusion, that indirect effects are far more important than direct ones, has been questioned because the definition of "indirect" paths (e.g., 1--> 2--> 5) seems to include many "direct" (but time-delayed) paths (e.g., 1--> 1--> 5)[46,48].

At this point, the matrix algebra used by Patten[46] and Levins[44] branch, although they support each other. Having presented the basic concepts of path analysis, we now consider some results of Levins[41] who summarizes overall interactions in the form of correlation matrices.

The links among a set of n variables, $X_i \ldots X_n$, come from differential equations describing the rates of change, dX_i/dt, of each variable as a function of all variables, $f_i(X_1, X_2, X_3, \ldots X_n)$. The link from X_j to X_i, given by $\alpha_{ij} = \delta f_i/\delta X_j$, is 0 if there is no connection from X_j to X_i; > 0 if there is a positive effect from X_i to X_j; < 0 if the effect of X_i diminishes X_j. The following definitions and rules[49] then permit the analysis of model systems.

a. A loop of length K is a simple, closed path from a variable to itself through k steps which visits each variable on the loop only once. The value of a loop is the product of the α_{ij} of its links, and the sign is the sign of that product. Feedback is defined as the effect of a variable on itself by way of intervening variables.

b. Mathematically, the feedback at level k, (F_k), in a system of n > k variables is defined by $F_k = \Sigma(-1)^{m+1}L(m,k)$. Feedback at level k is summed over all sets of the products of m disjunct loops that total k elements. Disjunct loops have no variables in common (L = loops).

c. Loops of length 0 have a value of +1 and $F_0 = -1$. This is an algebraic convenience.

d. A path P_{kj} (k) is a product of (k - 1) alpha values from X_j to X_i involving k variables, none of which are visited more than once. $P_{ii} = 1$.

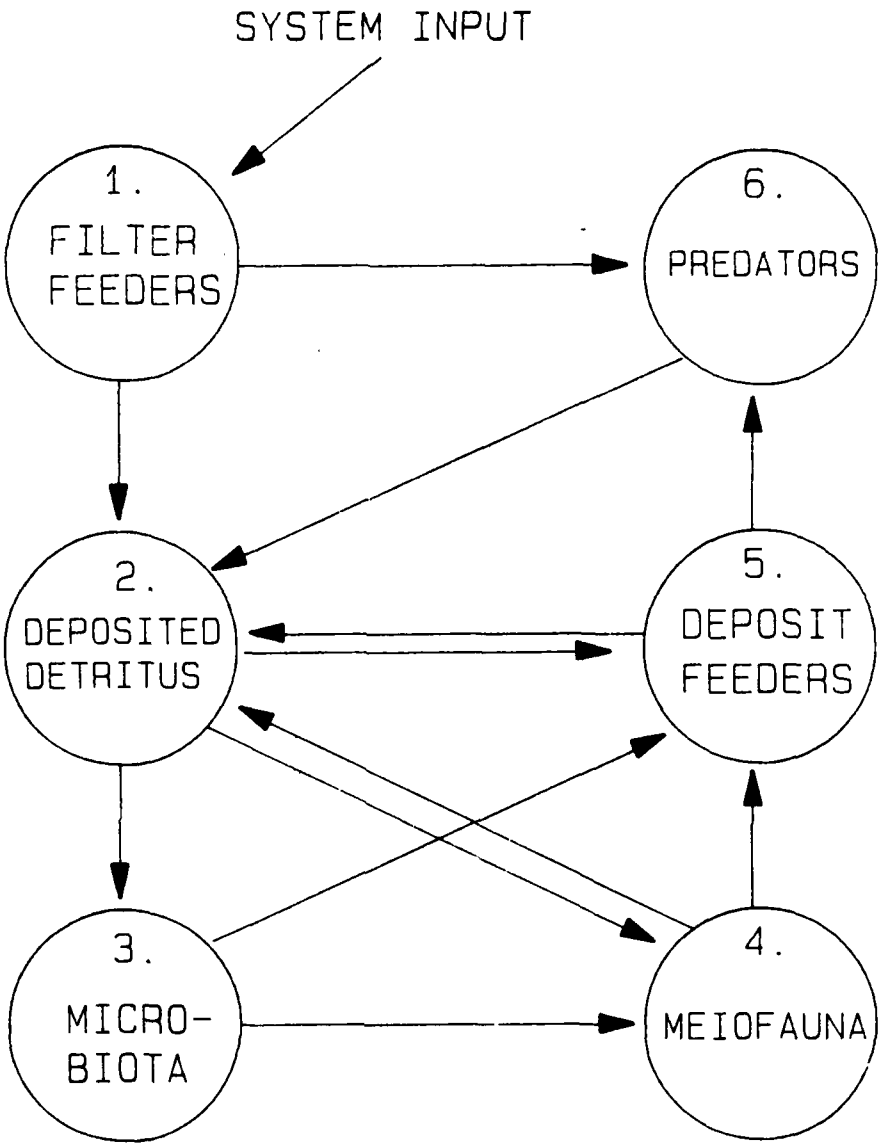


Figure 1. Oyster reef ecosystem energy flow compartment model. In addition to the paths shown, all compartments have a self-loop.

Table 7
Oyster Reef Model First-Order (Adjacency) Matrix[46,47]

To	From:	Compartments 1	2	3	4	5	6	Row Sum
1		1	0	0	0	0	0	1
2		1	1	0	1	1		5
3		0	1	1	0	0	0	2
4		0	1	1	1	0	0	3
5		0	1	1	1	1	0	4
6		1	0	0	0	1	1	3
Column sum		3	4	3	3	3	1	18

Table 8
Oyster Reef Model Second-Order Matrix for Paths[46,47]

To	From:	Compartments 1	2	3	4	5	6	Row Sum
1		1	0	0	0	0	0	1
2		3	3	2	3	3	2	16
3		1	2	1	1	1	1	7
4		1	3	2	2	1	1	10
5		1	4	3	3	2	1	14
6		2	1	1	1	2	1	8
Column sum		9	13	9	10	9	6	56

e. The complement of a path is the set of variables not on the path.

f. Let C_h be any of s parameters of the system, $f_i(X_1,X_2,X_3,\ldots,X_n; C_1,C_2,C_3,\ldots,TC_s.)$. Then the effect of a change in C_h on the equilibrium level of any variable (X_j) in the system obeys the following rule:

If C_h is a positive input to X_i, then its effect on X_j will have the sign of the sum of the products of each path from X_i to X_j, each multiplied by the feedback of its complement, and all divided by the feedback of the whole.

Several qualitative results follow.

a. Since $F_n < 0$, if the complementary subsystem of a path is stable, its feedback is also negative, and the equilibrium level of X_j in the system has the same sign as the path products if they are all the same.

b. If the complement has zero feedback for all paths, then the equilibrium level is independent of C_h.

c. If the complement of a path has positive feedback, the path has an effect of opposite sign to its own product.

d. The closer F_n approaches 0 (instability due to positive feedback equaling negative), the more sensitive all equilibrium values are to parameter change.

e. If a variable links with the other variables of a system in such a way as to contribute mostly negative feedback, $F_n > F_{n-1}$, the feedback of its complement, then it will be relatively insensitive to changes in its own input.

Consider a simple system consisting of one nutrient source, N, two herbivores H_1 and H_2, and a consumer, C, of one of these. Denoting a positive effect by --> and a negative effect by --O, the components are related by the graph in Figure 2. There may be many interactions, known and unknown, between these components: H_1 may inhibit H_2 by some toxin, C may stimulate the growth of H_1, etc. Since the impact cannot be predicted if the structure of the network is not known, we can examine possible impacts by representing alternative models as such signed digraphs. For each one, an effects matrix like Table 9 is constructed which indicates the direction of change of the variable listed above each column when the direct impact of a chemical enters the system as a positive input through the variable at the left of each row (Table 9). This analysis shows that, for many of the predictions, the detailed structure does not matter. The predictions that coincide under different models are robust whereas those differing between models are disjunct and can be used to decide among them.

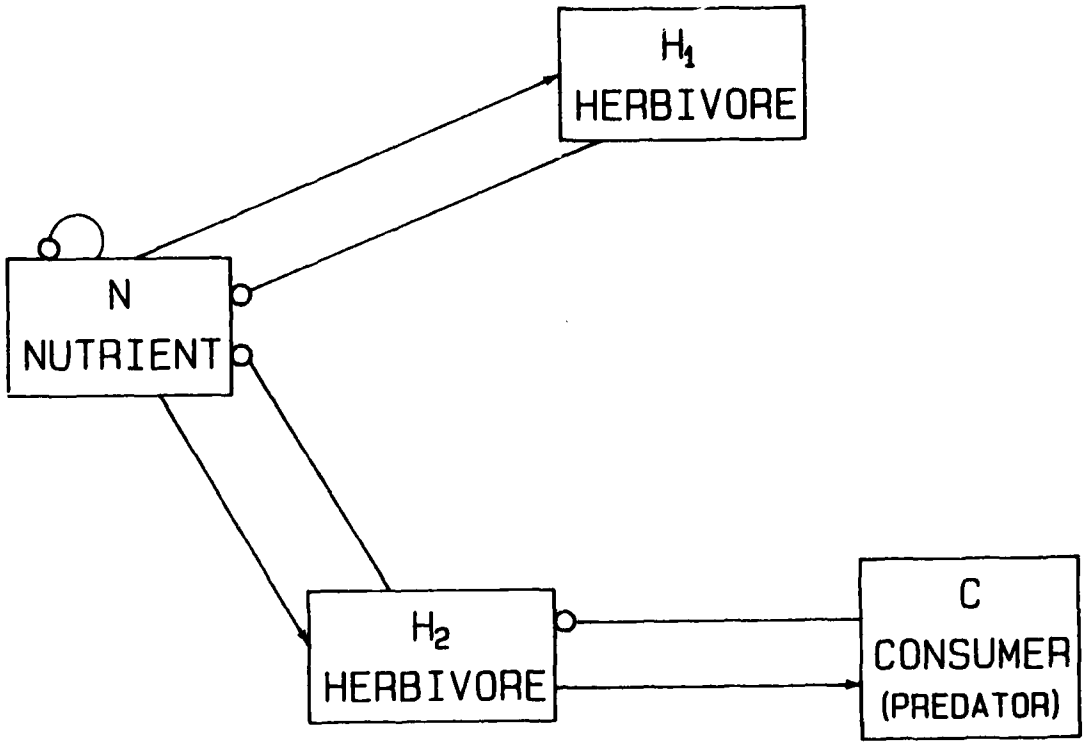


Figure 2. Digraph of a nutrient-driven simplified ecosystem.

Table 9
Matrix of Effects for the Signed Digraph of Figure 2

Impact Enters	Variables That Change			
	N	H_1	H_2	C
N	0	+	0	0
H_1	-	+	0	-
H_2	0	0	0	+
C	0	+	-	0

This table shows that an increase in the nutrient level, N, has no effect on itself (because it is self-damped). However, there is a direct increase in the numbers of H_1. Although H_2 is also directly increased by addition of nutrient, this increase is balanced by increased predation by the consumer. Because the consumer itself is not directly connected to the nutrient source, its numbers are unaffected. If H_1 is added to the system, the level of nutrient decreases. Although the total amount of nutrient available for the production of H_2 decreases, the graph shows that, because H_2 is not connected to H_1, the numbers of H_2 are unaffected. In order for this to occur, however, the numbers of consumers must decrease. The addition of H_2 to the system is compensated by increases in the numbers of consumers, so the levels of nutrient, H_1 and H_2, are unaffected. Finally, as consumers are added to the system, the numbers of H_2 are directly decreased. The release or additional availability of nutrients is accompanied by an increase in the numbers of H_1, so that there is no net change in the quantity of nutrient.

Consider a model (Figure 3) of a cultivated field community with a crop plant (P_1) population self-limited by crowding. The herbivore H_1 is the pest species that eats the plant. The specialized parasitoid, P_a, say a wasp, kills only H_1, while the generalized predatory insect, a spider, P_r, eats H_1 and H_2, a herbivore that feeds on other plants. The effect of adding pesticide, I, to this system is of interest. In the following discussion, the complements of a path are all remaining elements that are not part of that path. For example, the complements of the direct path [I, H_1] are P_1, P_a, and P_r--H_2--P_2.

An increase in the level of insecticide use, I, has a direct negative effect on P_a, P_r, and H_1. But the direct path [I,H_1] has zero feedback because the parasitoid is isolated. Therefore, this path has no effect. The other two paths [I, P_a, and H_1] are both positive, but only the first of these has a nonvanishing complement [P_1 and P_r--H_2--P_2]. Therefore, the final result of adding insecticide at a new, constant dosage is to increase the herbivore species and leave events further along the path P_r--H_2--P_2 unaltered. For example, if H_2 increases, this increases P_r, which eats more

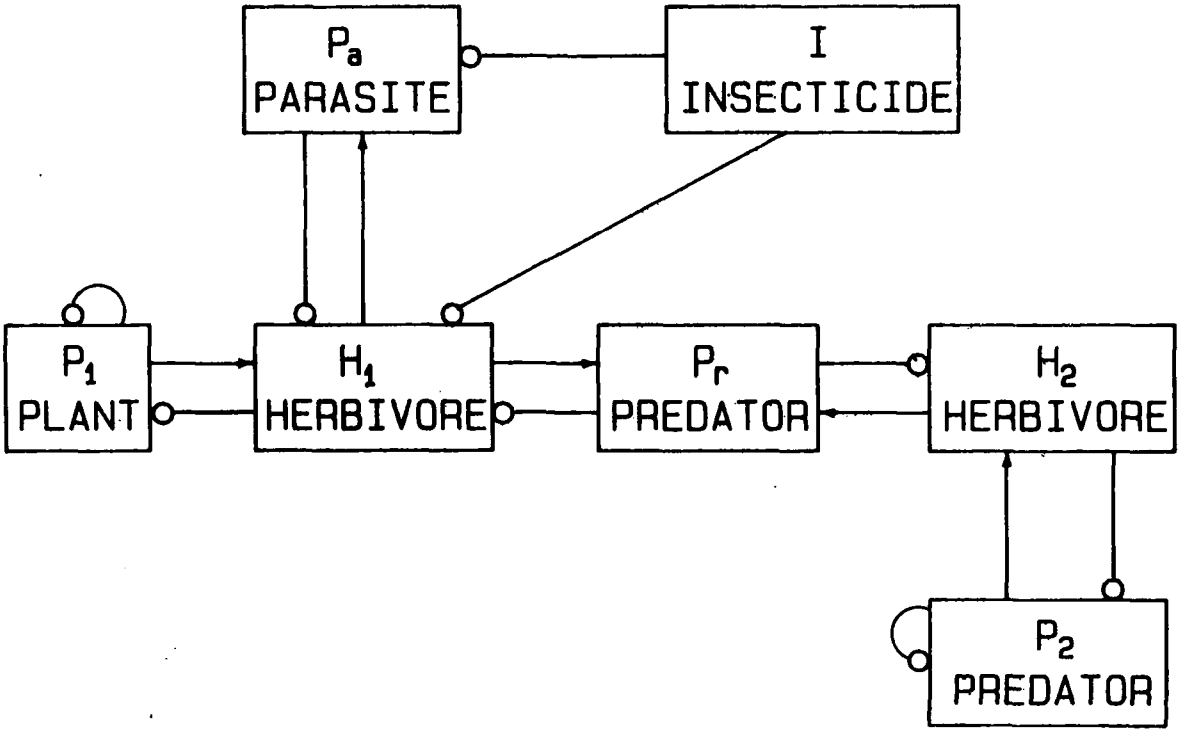


Figure 3. Digraph of a cultivated field under insecticide use.

H_1, reduces the food supply for P_a, and results in a shift in the cause of death of H_1 but not its numbers. The only way to affect H_1 is through P_a. This argument, of course, holds only if the system is maintained near equilibrium; high enough dosages of pesticide can destroy the equilibrium and wipe out the species.

3.4.2 Using These Models.

At this phase of development, the ecologist's task is to qualitatively describe the ecosystem under stress from simulants as concisely as possible. When feasible, these descriptions should be in terms of one or more of the models given by Levins[45] so that his effects matrices can be used directly. When this is not possible, his original papers will have to be consulted for the methods needed to construct the effects matrices. It is unlikely that the ecologist will be able to implement these methods without the assistance of a good mathematician or modeler.

Levins' models offer two useful possibilities. From the pattern of population swings, it may be possible to identify the entry point(s) of impacts and/or the larger correlations and their signs between species populations. Alternatively, when population interactions are fairly well known in advance, the damage points of products can be projected even if damage is indirect.

4. DECISION TREE FOR PRODUCTS

We seek a protocol for projecting effects on unspecified ecological systems/processes and their components in the absence of knowledge of which effects on which aspects of ecosystems are to be considered significant. The open-ended nature of the target and effects sets necessarily expand the number of substances surviving the elimination series of models (e.g., battery) and the number of possible effects found for each. Consequently, our models' output should be evaluated in the light of output from a model designed to identify critical features of those ecological systems at the Army bases where simulants and decontaminants are used. Such an evaluation can be expected to result in rank shifts and the elimination of some simulants and decontaminants from the ecohazard list.

The method of Cramer, et al.[50], that has stood the test of 10 years of review, is applicable to this proposal for setting priorities only, not for making final judgments of hazard or risk. Such judgments can be made only by a group of experts and must take into account the following: (1) best professional consensus judgment of toxicological hazard; (2) duration of use-experience; (3) quantity used; (4) number of people exposed; (5) probability of a toxic response at that exposure; (6) expected severity of a toxic response at that exposure; (7) availability of less hazardous materials for the same purpose; and (8) formal evaluations by other qualified experts such as the FDA in approving food additives. In the latter case, however, caution must be used because of the fact that such evaluations may be for totally different conditions of exposure. It is imperative to consider use conditions when making hazard assessments.

4.1 Research Data.

Risk can be assessed only if an experimental data base on the product and its contaminants is available; even then, there is a degree of uncertainty. The process starts by scanning the population of materials to be considered by the system. This population consists of materials in use and proposed for use. Information required for hazard identification includes product composition, proposed use(s), known and possible contaminants, physical and chemical properties, and likely toxicological effects to any species. Product test data include experimental data from extraction tests on a product, specific chemical contaminant levels of group parameter levels (total extractable solids, etc.), and all available toxicological data. Some of these data will be available from the toxicology or epidemiology literature and some from the manufacturer's process data and tests. In many cases, however, such data are unavailable. Since very little toxicological data exist for the majority of unique chemical structures[15], additional tests may be required. The formal evaluation process begins with a dossier (Table 10).

Reliable tests for genotoxicity will require an understanding of the biochemical and physical-chemical mechanisms operating in the test system and under the real conditions of ecological systems exposure which we do not yet have. Hence, while the Decision Tree proposed here will allow for systematic decision-making, it will not eliminate all risks associated with the use of products. It is not possible to reach the goal of zero risk other than in the case where exposure is zero--a situation that is illusive and vague in itself.

4.2 Exposure Assessment.

This activity involves the determination of likely exposures by target species. The data for this assessment are the results of Product Tests devised to determine if a given product will release any contaminant of environmental concern. Three experimental parameters are of interest: the identity of contaminants released, contaminant concentrations, and release schedules. Thus, data must be developed for each category of product and used precisely to evaluate each product.

Training use scenarios for products are critical elements of exposure assessment. Materials used at prepared ranges or indoors should present no ecohazard. Simulants sprayed on troops or equipment and the decontaminants then used for "hasty" cleanup (except for personal decontamination kit materials) are more likely to be ecotoxic, for they may be deposited directly in natural communities.

Currently used products judged to have a high, and perhaps medium, hazard potential would be subjected to the more detailed assessment envisaged in the Decision Tree for new products. Products passing this review, and which already have health-safety certifications for related uses, would be certified for use. Products passing the review but not already having health certifications through other mechanisms would also be approved. Products already in use that do not meet these demonstrations and new products require a formal hazard determination. For most products, it should be possible to make this determination from the submission documentation. The types of documentation acceptable for this purpose are structure-activity studies on raw materials

Table 10
Synopsis of Major Components of a Dossier (Modified From Reference[15])

*1. Information Desirable for Initial Hazard Review
- Chemical name, CAS Registry number, and intended-use category
- Physiochemical properties, manufacturing processes, production, uses, chemical fate, and exposure potential
  - Physiochemical properties
  - State(s)
  - Solubility (fat, water, e.g., octanol-water partition coefficient)
  - Chemical uniqueness
  - Manufacturing process
  - Production level
  - Consumption level
  - Uses (intended and other known uses)
- Roadblocks to acquiring above information
- Level of concern based on all the above information

*2. Product Test Data
- Statement of chemical fate (data aimed at determining actual or expected exposure)
- Exposure potential determined from extraction and other tests.
- Production of residue by leaching; degradation products and products generated by reaction with chlorine; components causing tastes or odor:
  - In intended-use setting (dose, duration, frequency, route, number exposed)
  - In occupational setting (dose, duration, frequency, route, number exposed)
  - In environmental setting (dose, duration, frequency, route, number exposed)
- Roadblocks to acquiring above information
- Level of concern based on all above information

*3. Synopsis of the Health-Effects Data Base
- Summary of the toxicity data base
  - Human
  - Animal
- Structure-activity and dose-response relationships
- Toxicologic uniqueness (e.g., reversibility of effect)
  - For required tests
  - For tests not required
- Roadblocks to acquiring toxicity information
- Level of concern based on all above information

**4. Adequacy of the Data Base
- Analysis of individual toxicity studies
- Analysis of the complete toxicity data base
- Analysis of the complete data base

*These data to be supplied by manufacturer.
**This evaluation to be made by Certification Board.

and products; laboratory studies of various types, such as compositional analysis; leaching and residue analyses[51,52]; and toxicological data from screening or definitive assays on raw materials and/or residues from leaching tests.

4.3 Decision Tree Method for Improving Rank Matrix.

We use a decision tree to develop parsimonious models that can be used systematically to assess probable impacts of products (simulants, decontaminants, carriers, use formulations, and breakdown products) on ecological systems. We move from models with limited predictive ability and data requirements to ones less restrictive in both inputs and outputs. Decision-tree models employ sequentially arranged binary choice questions to data for elimination of action pathways from further consideration. They require a complete set of the assumed input data and well defined targets. While the approach is general, we are specifically concerned with listed simulants, decontaminants, and carriers. Identified breakdown products can also be evaluated.

The tree initially determines if the compound is designed for use, or is used, as a biocide. We require that toxicological data for biocides come from experiment whereas data for nonbiocides can be estimated using supplied algorithms. Generally, these algorithms require only physical and chemical properties such as water solubility and vapor pressure.

4.4 The Decision Tree.

a. Is the compound part of the personal decontamination packet, used only indoors, or used only at prepared chemical training ranges?

Yes--assign ecological system hazard rank 0. No further consideration necessary for small quantities properly disposed of, the usual training conditions.

b. Is the compound a biocide (other than bactericide)?
Yes--use laboratory toxicity data
No---use laboratory toxicity data or estimated values.

c. Is the compound acutely toxic to native fauna?
Several approaches are suggested for assessing the acute toxicity of chemicals to mammals. In method (1), one-dose acute toxicity data for a test species is converted to human exposure equivalents[55] (in mg/kg). In method (2), an acute toxicity value for one laboratory species is used to estimate a hazard-equivalent exposure for the target species[55,56]. This value is compared with the levels estimated from PPLV models to be present in the field. Method (3) is algorithms for converting acute toxicity data between species of fishes.[9,12,22,53,54] Method (4) is algorithms for converting acute toxicity data between rodents and fishes and vice-versa[54]. None of these methods will work if the compound is a cholinesterase inhibitor, such as an organophosphate or carbamate pesticide. Cholinesterase inhibitors are identified as such in the

environmental data sheet reports[2,3]. Method number (5) is SAR models[1,30,33-36]

(1) Compute the acute toxic exposure for a standard target organism, here taken as man, D_{MAN}, from the toxicity for the known organisn, D_{ANIMAL}. The allometric relationship used is based on the relationship between metabolism and body weight, W:

$$D_{MAN} = D_{ANIMAL} \quad (W_{ANIMAL}/W_{MAN})^{0.25} \qquad (7)$$

Values of the scaling factor, $(W_{ANIMAL}/W_{MAN})^{0.25}$, between man and other mammals are given in Table 11[56] and acute toxicity rank scores in Table 12.

(2) Table 13 gives scaling factors calculated for all animal weight combinations. In many instances, estimates of acute toxicity exposure concentrations are needed for a particular target species for which there are no toxicological data. Table 13 provides conversion factors that can be used to scale available toxicity information to estimate the acute toxic exposure level for the target species. The estimated toxic exposure value is compared with levels estimated from the PPLV models for the primary environmental media.[4] If the levels in the environment exceed the estimated toxic dose, a rank of 6 is assigned; otherwise, the rank is 3.

(3) Several models for estimating acute toxicity in bluegills and fathead minnows are given in Table 14[54]. Separate models are provided for general organics and noncholinesterase inhibitor pesticides. For example, the LD50 for fathead minnows estimated from bluegill LD50 of 10 ppm (first data line of Table 14) for a chemical of molecular weight 200 is:

$$\text{Fathead LD50} = 0.94 + 0.287 \log (10/200) = 0.95$$

(4) Models can also be used to estimate acute toxicity in fish and rats from each other as shown in Table 15.[54]

(5) SAR models can be used to estimate acute toxicity when it cannot be measured directly. Estimates can be made for standard laboratory species using one or more of several models such as octanol-water partition coefficient and SAR relationships. Allometric adjustments can generally be made for mammals.

Data now are available on the chemical, physical, and toxicological characteristics of a large number of chemicals, as is information on their transport, transformation, and environmental fate. Some correlations now can be made between the structure of certain chemicals and the properties they exhibit, a process called "structure activity analysis." Under favorable conditions, these correlations allow rough predictions to be made about the characteristics of chemicals that have not been studied, reducing the amount of testing necessary to make at least an initial judgment of the risks they may pose. Structure-activity relationships can be used to delineate the

general properties of interest, but they are not generally a substitute for actual measurements[12]. As noted in the Federal Register[57]:

> The use of structure-activity relationships may allow preliminary exposure estimates to be made even when there is a scarcity of data on a specific chemical. Structure-activity analysis is a relatively new field, and the available tools are still crude. The user must exercise scientific judgment in interpreting the results, because substantial work remains to be done in refining and validating these techniques.

d. Is the compound likely to display chronic toxicity? Following the reasoning used in developing safety factors for PPLV analysis,[4] we propose that the chronic toxicity exposure for fauna be estimated as:

0.01 (human-standardized acute exposure level).

e. Is the compound genotoxic? Genotoxic effects may be important if the natural community contains species of low reproductive potential, such as large mammals. Without presenting the detailed rationale, studies unrelated to this contract have shown that there is a high statistical correlation between mutagenicity and nonmutagenicity in genetically engineered strains of Salmonella bacteria and between carcinogenicity and noncarcinogenicity in rodents. Using an approach similar to that for acute toxicity, we propose assigning the following ranks (Table 16) to the dose/plate (μg), causing a doubling of the background response.

f. Is the compound phytotoxic? This question cannot be answered for most products at this time. We have not found models for phytotoxicity that would be useful for chemical agent products. Modeling in this area appears to be limited to descriptions of toxicity variations within narrow families of pesticides. No general rules have emerged, although a 2300-page compilation of chemicals tested for phytotoxicity is available.[59] A search for more recent data on phytotoxic responses to any of 21 simulants and decontaminants by Oak Ridge National Laboratory's Toxicology Information Response Center revealed no new information.

g. Is the compound likely to bioaccumulate? If the log of the octanol-water partition coefficient is greater than 3.5, assume a bioaccumulation risk, unless the material is rapidly degraded[12,25].

h. Is soil accumulation and consequent chronic exposure a concern? Soil accumulation levels can be estimated from octanol-water partition and breakdown data[12,25].

i. Does the product or its hydrolysate have an extreme pH (outside the range of 5.0 to 9)? Such materials, when used in amounts contemplated for training, present an acute, but not chronic, threat to biota.

Example 1: Sodium Benzenesulfonamide, $C_6H_5SO_2NClNa$.

15 mg/kg--pulmonary edema in rat
Water-soluble
Does not decompose
No mutagenicity data.

a. Exposure assessment: The most likely route of ecosystem exposure is ingestion by mammals of disposed towelettes. No significant exposure of plants or fish is possible. Exposure rank = 0.

b. Is the compound a biocide: No. Compound is a bactericide.

c. Conversion to standard organism: If the observed toxic effect occurs in man, the expected scaling factor is 0.251. Hence,

$$D_{MAN} = 0.251\ (15\ mg/kg) = 3.77\ mg/kg$$

Rank: 6 supertoxic

Hazard ranking: Toxicity rank x exposure rank = 0.

Conclusion: it is unlikely that this compound poses a toxic hazard to mammals, fish, or plants. However, exposures to mammals can be eliminated by rewrapping used towelettes in foil and discarding them in appropriate containers.

Example 2: n-Butyl Mercaptan, $CH_3(CH_2)_3SH$.

Rat oral LD50--1500 mg/kg
Rat inhalation LC50--4020 ppm/4 hr
Mouse inhalation LC50--2500 ppm/4 hr
Rabbit eye irritation--83 mg
Nonmutagenic
Nonteratogenic
Cholinesterase antagonist
Slightly soluble in water
Highly volatile.

a. Exposure assessment: ecohazard arises from spray use in unit field training. Because of its high volatility and low mammalian toxicity, other exposure routes can be ignored.

b. Is the compound a biocide? No. The compound has weak anticholinesterase activity.

c. The compound's high volatility and low water solubility protect fish from it. Expected native mammal acute toxicity is nominally medium, but native mammals tend to avoid acute exposure during training exercises. Birds are more likely to experience exposure. Converting of the 1500 mg/kg acute rat LD50 to our standard organism, man: $D_{man} = 0.251\ (1500\ mg/kg) = 376.5$ mg/kg, this corresponds to an NRC acute toxicity rank[6] of medium (see page 11).

d. Chronic native mammal toxicity is zero because the high volatility prevents chronic exposure.

e. The compound is not genotoxic.

f. Phytotoxicity: mercaptans interfere with plant energy assimilation, which is of concern in agriculture but not of importance for natural systems that tend to be limited by other factors.

g. Due to high volatility, n-butyl mercaptan does not bioaccumulate.

h. Soil accumulation is not of concern because of high volatility.

i. pH: very weak acid.

Hazard ranking: exposure rank is composed of a high area coverage and an estimated low exposure, since only vapor and direct contact are available as exposure routes. We will treat this conservatively as the low exposure, high EBCA, and medium toxicity case in Table 1, yielding a rank of 5.

Conclusion: n-butyl mercaptan presents a moderate danger to ecosystems, largely because of expected direct application to natural communities. Birds are probably the only community elements at significant risk.

Table 11
Scaling Factors Between Exposure in Animals and Humans[56]

Organism	Body Weight (g)	Scaling Factor
Mouse	20	0.141
Squirrel	100	0.211
Rat 200	0.251	
Guinea pig	400	0.299
Rabbit	1500	0.416
Cat 2000	0.447	
Monkey	4000	0.532
Infant (0-1 yr)	5000	0.562
Dog 12000	0.700	
Child (1-13 yr)	20000	0.795
Woman, Deer	50000	1.000
Goat, sheep, pig	60000	1.047
Man 70000	1.088	
Cattle, horse, donkey	500000	1.778

Table 12
Classification of Toxicants Into Categories of Relative Toxicity[55]

Toxicity Rank	Commonly Used Term	Probable Human Lethal Exposure 70-kg (150-lb) Man	
6	Supertoxic	<5 mg/kg	A taste; <7 drops
5	Extremely toxic	5-50 mg/kg	7 drops-1 tsp
4	Very toxic	50-500 mg/kg	1 tsp-1 ounce
3	Moderately toxic	0.5-5 g/kg	1 oz-1 pint or pound
2	Slightly toxic	5-15 g/kg	1 pint-1 quart
1	Practically nontoxic	>15 g/kg	>1 quart

Table 13
Conversion Factors for Various Species Based on Body Weight

		Target Species* (Denominator)													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Weight (Kg)		20	100	200	400	1500	2000	4000	5000	12000	20000	50000	60000	70000	500000
Known Species (Numerator)															
1	20	1.000	0.669	0.562	0.473	0.340	0.316	0.266	0.251	0.202	0.178	0.141	0.135	0.130	0.080
2	100	1.495	1.000	0.841	0.707	0.508	0.473	0.398	0.376	0.302	0.266	0.211	0 202	0.194	0.119
3	200	1.778	1.189	1.000	0.841	0.604	0.562	0.473	0.447	0.359	0.316	0.251	0 240	0.231	0.141
4	400	2.115	1.414	1.189	1.000	0.719	0.669	0.562	0.532	0.427	0.376	0.299	0.286	0.275	0.168
5	1500	2.943	1.968	1.655	1.392	1.000	0.931	0.783	0.740	0.595	0.523	0.416	0.398	0.383	0.234
6	2000	3.162	2.115	1.778	1.495	1.075	1.000	0.841	0.795	0.639	0.562	0.447	0.427	0.411	0.251
7	4000	3.761	2.515	2.115	1.778	1.278	1.189	1.000	0.946	0.760	0.669	0.532	0.508	0.489	0.299
8	5000	3.976	2.659	2.236	1.880	1.351	1.257	1.057	1.000	0.803	0.707	0.562	0.537	0.517	0.316
9	12000	4.949	3.310	2.783	2.340	1.682	1.565	1.316	1.245	1.000	0.880	0.700	0.669	0.643	0.394
10	20000	5.623	3.761	3.162	2.659	1.911	1.778	1.495	1.414	1.136	1.000	0.795	0.760	0.731	0.447
11	50000	7.071	4.729	3.976	3.344	2.403	2.236	1.880	1.778	1.429	1.257	1.000	0.955	0.919	0.562
12	60000	7.401	4.949	4.162	3.500	2.515	2.340	1.968	1.861	1.495	1.316	1.047	1.000	0.962	0.589
13	70000	7.692	5.144	4.325	3.637	2.614	2.432	2.045	1.934	1.554	1.368	1.088	1.039	1.000	0.612
14	500000	12.574	8.409	7.071	5.946	4.273	3.976	3.344	3.162	2.541	2.236	1.778	1.699	1.635	1.000

*Organisms:

1	Mouse	8	Infant (0-1 yr)
2	Squirrel	9	Dog
3	Rat	10	Child (1-13 yr)
4	Guinea pig	11	Woman, Deer
5	Rabbit	12	Goat, sheep, pig
6	Cat	13	Man
7	Monkey	14	Cattle, horse, donkey

Table 14
Regression Models for Bluegill and Fathead Minnow

X Variable*	Y Variable	r	B	a
A. General Organic Compounds				
Bluegill	Fathead	0.94	0.94	0.287
Fathead	Bluegill	0.95	0.94	0.174
B. Pesticides (not including organophosphates or carbamates)				
Bluegill	Fathead	0.96	1.11	0.612
Fathead	Bluegill	0.91	0.79	0.530
C. General Organic Compounds and Pesticides Combined				
Bluegill	Fathead	0.95	0.96	0.395
Fathead	Bluegill	0.90	0.88	-0.300
D. Pesticides--Janardan, et al.[54] compared with Kenaga[9]**				
Fathead	Bluegill	0.91	0.78	0.024 (Janardan)
Fathead	Bluegill	0.85	0.82	0.01 (Kenaga n=21)
Fathead	Bluegill	0.87	0.86	0.066 (Kenaga n=7)

*Log (LD50/w) or log (LC50/w), log = logarithm base 10,
w = molecular weight.
**Log (LD50) or log (LC50).

Table 15
Regression Models for Rats Versus Fish[54]

X Variable*	Y Variable	r	B	a
A. Organic Compounds				
Rat	Fathead	0.77	1.37	0.799
Fathead	Rat	0.63	0.35	-0.161
Rat	Bluegill	0.71	1.21	0.539
Bluegill	Rat	0.74	0.43	-0.056
B. Organic Compounds and Chlorinated Pesticides Combined**				
Rat (female)	Fathead	0.83	1.53	0.689
Fathead	Rat (female)	0.67	0.36	-0.259
Rat (male)	Fathead	0.65	1.15	0.820
Fathead	Rat (male)	0.58	0.33	-0.340
Rat (female)	Bluegill	0.68	1.04	0.492
Bluegill	Rat (female)	0.75	0.49	-0.313
Rat (male)	Bluegill	0.66	1.04	0.428
Bluegill	Rat (male)	0.73	0.47	-0.272
C. Chlorinated Pesticides				
Rat (male)	Bluegill	0.88	1.45	-0.639
Bluegill	Rat (male)	0.76	0.46	0.125
Rat (female)	Bluegill	0.999	1.51	-0.521
Bluegill	Rat (female)	0.92	0.66	0.345
Rat (male)	Fathead	0.98	1.70	-0.326
Fathead	Rat (male)	0.999	0.59	0.192
Rat (female)	Fathead	0.96	1.29	-0.490
Fathead	Rat (female)	0.999	0.28	0.380

*Log (LD50/w) or log (LC50/w), log = base 10, w = molecular weight.
**Male and female refer to pesticide data.

Table 16
Percentage Points at the 95 Percent Confidence Level of the Cumulative Distribution of Doubling Doses of Mutagens in the *Salmonella* Plate Assay

Percentage of population included[58]	Doubling dose (μg/plate)[58]	Proposed Hazard Classification	Rank
20	0.20	Super mutagenic	6
30	0.36	to very mutagenic	
40	0.65		
Class boundary	1.0	Moderately mutagenic	5
50	1.7		
60	3.0		
Class boundary	10.0	Slightly mutagenic	4
70	13.3		
80	46.2		
90	185		3
Class boundary	1000	Practically	2
95	1500	nonmutagenic	
98	4750		1

5. RECOMMENDATIONS

a. Minimal environmental data sheet information should include phytotoxicity data and directly measured octanol-water partition coefficients.

b. Environmental data sheets should be prepared for breakdown products likely to occur in substantial amounts.

c. Because ecotoxicology is so primitive, it is unwise to rely on any of the projection techniques given. We urge "experimentation"--observation during training substance use--on the large system. Training areas are small and embedded in larger areas that can supply time-controlled baseline information and disseminules of species eliminated by simulant and decontaminant use. Good ecological inventory and chemical sampling now on training areas and their environs would be a spatial and historical baseline against which degradation and associated chemical residues could later be evaluated. Such a scheme should be implemented because there is great uncertainty in our uninformed assessment of likely damage. In general, the current state of ecotoxicology knowledge is such that projection should be viewed with suspicion. Projection is justified only when ecosystem measurements cannot be made. This is well expressed in a recent review of the limitations of laboratory bioassays[60]:

> We must be aware of the limitations of any prediction scheme; we must recognize the need for continuous monitoring of ecosystem properties to test our predictions and provide a basis for management decisions.

For economy's sake, those ecosystem variables should be limited in number. Because our knowledge of interrelationships in particular systems is so poor, monitor variables should address as directly as possible those ecosystem features important to the mission or legal requirements.

d. Priority for completion of dossiers should be those materials applied to troops and equipment in unit field exercises.

LITERATURE CITED

1. Benford, P. J. and [illegible] I. Wenz, Environmental Overview of Decontaminating Agents, Draft Report (U.S. Army Materiel Command, 1985).

2. Bennett, S. R., J. M. Bane, P. J. Benford and R. L. Pyatt, Environmental Hazards of Chemical Agent Simulants, Technical Report CRDC-TR-84055 (U.S. Army Armament, Munitions and Chemical Command, Aberdeen Proving Ground, MD, 1984).

3. Howard, P. H., G. W. Sage, J. P. Robinson and J. Jackson, Environmental Fate Assessments of Chemical Agent Simulants, Technical Report CRDC-CR-86016 (U.S. Army Armament, Munitions and Chemical Command, Aberdeen Proving Ground, MD, 1986).

4. Small, M. J., The Preliminary Pollutant Limit Value Approach: Procedures and Data Base, Technical Report 8210, ADA150767 (U.S. Army Medical Bioengineering Research and Development Laboratory, Fort Detrick, MD, 1984).

5. U.S. Environmental Protection Agency, "National Oil and Hazardous Substances Contingency Plan," Fed. Regist., Vol 47, No. 137 (16 July 1982), pp 31180-31243.

6. National Research Council, Risk Assessment in the Federal Government: Managing the Process (National Academy Press, 1983).

7. Stevens, K. R. and M. A. Gallo, "Practical Considerations in the Conduct of Chronic Toxicity Studies," In A. W. Hayes, ed. Principles and Methods of Toxicology (Raven Press, 1982), pp 53-79.

8. Petrocelli, S. R., "Chronic Toxicity Tests," In G. M. Rand and S. R. Petrocelli, eds., Fundamentals of Aquatic Toxicology (Hemisphere Pub. Corp., 1985), chapter 4.

9. Kenaga, E. E., "Test Organisms and Methods Useful for Early Assessment of Acute Toxicity of Chemicals," Environ. Sci. Technol., Vol 12 (1978), pp 1322-1329.

10. Sheehan, P. J., "Effects on Individuals and Populations," In P. J. Sheehan, D. R. Miller, G. C. Butler and Ph. Bordeau, eds., Effects of Pollutants at the Ecosystem Level (SCOPE 22. John Wiley and Sons, 1984), p 38.

11. Waldichuck, M., "Review of the Problems," Phil. Trans. R. Soc. London. B. Biol. Sci., Vol 286 (1979), pp 399-424.

12. Lyman, W. J., W. F. Reehl and D. H Rosenblatt, Handbook of Chemical Property Estimation Methods (McGraw-Hill Book Co., 1982).

13. Schaeffer, D. J. and K. G. Janardan, "Designing Batteries of Tests with Largest Inter-tier Correlation," Ecotox. Environmen. Safety. (1987 in press).

14. Lave, L. B., G. S. Omenn, K. D. Heffernan, and G. Dranoff, "Analysis of the Cost-Effectiveness of Tier-Testing for Potential Carcinogens," a paper presented at the First World Conference of the American Association of Toxicologists, 29 May 1982, and summarized in Toxicity Testing, Strategies to Determine Needs and Priorities (National Academy Press, 1984), pp 275-282.

15. National Research Council, Toxicity Testing: Strategies to Determine Needs and Priorities (National Academy Press, 1984).

16. Odum, E. P., "The Strategy of Ecosystem Development," Science, Vol 164 (1969), pp 262-270.

17. Holling, C. S., "Resilience and Stability of Ecological Systems," Ann. Rev. Ecol. Syst., Vol 4 (1973), pp 1-24.

18. Westman, W. E., "Measuring the Inertia and Resilience of Ecosystems," Bioscience, Vol 28 (1978), pp 705-710.

19. Ferguson, J., "The Use of Chemical Potentials as Indices of Toxicity," Proc. R. Soc. London, Vol 127B (1937), pp 387-404.

20. Mullins, L. J., "Some Physical Mechanisms of Narcosis," Chem. Rev., Vol 54 (1954), pp 289-323.

21. Veith, G. D., State-of-the-Art Report on Structure-Activity Methods Development, EPA-600/3-81-029, NTIS PB81-187239 (USEPA, 1981).

22. Kenaga, E. E. and C. A. I. Goring, "Relationship Between Water Solubility, Soil Sorption, Octanol-Water Partitioning, and Concentration of Chemicals in Biota," In J. G. Eaton, P. R. Parrish and A. C. Hendricks, eds., Aquatic Toxicology, ASTM STP 707 (ASTM, Philadelphia, 1980), pp 78-115.

23. Yalkowsky, S. H. and G. L. Flynn, "Transport of Alkyl Homologs Across Synthetic and Biological Membranes--A New Model for Chain Length-Activity Relationships," J. Pharm. Sci., Vol 62 (1973), pp 210-217.

24. Kubinyi, H., "Non-Linear Dependence of Biological Activity on Hyrophobic Character: The Bilinear Model," Farmaco (Pavia) Ed. Sci., Vol 34 (1979), pp 248-276.

25. Gillett, J. W., "A Comprehensive Prebiological Screen for Ecotoxicologic Effects," Environ. Toxicol. Chem., Vol 2 (1983), pp 463-476.

26. Chapara, S. C. and K. H. Reckhow, "Engineering Approaches for Lake Management," Vol 2, Mechanistic Modeling (Butterworth, 1983).

27. Smith, J. H., W. R. Mabey, N. Bohonos, B. R. Holt, S. S. Lee, T. W. Ghou, D. C. Bomberger, and T. Mill, Environmental Pathways of Selected Chemicals in Freshwater Systems, Part I: Background and Experimental Procedures, EPA-600/7-77-113 (USEPA, 1977).

28. Burns, L. A., D. M. Cline, and R. R. Lassiter, Exposure Analysis Modeling Systems (EXAMS) Users Manual and Systems Documentation Report, EPA-800/3-82-023 (USEPA, 1982).

29. Hansch, C., G. J. Hatheway, F. R. Quinn and N. Greenberg, "Antitumor 1-(X-aryl)-3,3-dialkyltriazines. 2. On the Role of Correlation Analysis in Decision Making in Drug Modification. Toxicity Quantitative Structure-Activity Relationships of 1-(X-phenyl)-3,3-dialkyltriazines in Mice," J. Med. Chem., Vol 21, No. 6 (1978), pp 574-577.

30. Venger, B. H., C. Hansch, G. J. Hatheway and Y. U. Amrein, "Ames Test of 1-(X-phenyl)-3,3-dialkyltriazines: A Quantitative Structure-Activity Study," J. Med. Chem., Vol 22, No. 5 (1979), pp 473-476.

31. Purcell, W. P., "Quantitative Structure-Mutagenicity Relationships," in L. Goldberg, ed., Structure-Activity Correlation as a Predictive Tool in Toxicology: Fundamentals, Methods and Applications (McGraw-Hill, 1981), pp 221-228.

32. Hatheway, G. J., C. Hansch, K. H. Kim, S. R. Milstein, C. L. Schmidt, R. N. Smith, and F. R. Quinn, "Antitumor 1-(X-aryl)-3,3-dialkyltriazines: 1. Quantitative Structure-Activity Relationships vs. L1210 Leukemia in Mice," J. Med. Chem., Vol 21, No. 6 (1978), pp 563-574.

33. Enslein, K. and P. N. Craig, "A Toxicity Estimation Model," J. Environ. Pathol. Toxicol., Vol 2 (1978), pp 115-132.

34. Enslein, K., T. R. Lander, M. E. Tomb and W. G. Landis, "Mutagenicity (Ames): A Structure-Activity Model," Teratog. Carcinog. Mutagen., Vol 3 (1983), pp 503-513.

35. Enslein, K., and P. N. Craig, "Carcinogenesis: A Predictive Structure-Activity Model," J. Toxicol. Environ. Health, Vol 10 (1982), pp 521-530

36. Enslein, K., T. R. Lander and M. E. Tomb, "Comparison of Teratogenesis, Carcinogenesis and Mutagenesis in a Data Base," DHI Toxicol. Newsletter, No. 1 (1983), pp 2-3.

37. Spann, M. L., K. C. Chu, W. T. Wipke and G. Ouchi, "Use of Computerized Methods to Predict Metabolic Pathways and Metabolites," J. Environ. Pathol. Toxicol., Vol 2 (1978), pp 123-131.

38. Rapport, D. J., H. A. Regier and T. C. Hutchinson, "Ecosystem Behavior Under Stress," Am. Nature, Vol 125, No. 5 (1985), pp 617-640.

39. Hakanson, L., "Aquatic Contamination and Ecological Risk," Water Res., Vol 18, No. 9 (1984), pp 1107-1118.

40. O'Neill, R. V., S. M. Bartell and R. H. Gardner, "Patterns of Toxicological Effects in Ecosystems: A Modeling Study," Environ. Toxicol. Chem., Vol 2 (1983), pp 451-461.

41. Levins, R., "The Qualitative Analysis of Partially Specified Systems," Ann. N.Y. Acad. Sci., Vol 231 (1973), pp 123-138.

42. Levins, R., "Problems of Signed Digraphs in Ecological Theory," In S. A. Levin, ed., Ecosystem Analysis and Prediction (Soc. Indust. AppL. Math., 1975a), pp 264-277.

43. Levins, R., "Evolution in Communities Near Equilibrium," M. Cody and J. Diamond, eds., Ecology and Evolution of Communities (Belknap Press, 1975), pp 16-50.

44. Levins, R., "Coexistence in a Variable Environment," Am. Nature, Vol 114, No. 6 (1979), pp 765-783.

45. Levins, R., "Ecosystem Properties Relevant to Ecotoxicology," In Testing for Effects of Chemicals on Ecosystems, Appendix B (National Acadamy Press, 1981).

46. Patten, B. C., "On the Quantitative Dominance of Indirect Effects in Ecosystems," In Third International Conference on State-of-the-Art in Ecological Modeling (Colorado State Univ., 1982), pp 27-37.

47. Dame, R. F. and B. C. Patten, "Analysis of Energy Flows in an Intertidal Oyster Reef," Mar. Ecol. Prog. Ser., Vol 5 (1981), pp 115-124.

48. Wiegert, R. G. and J. Kozlowski, "Indirect Causality in Ecosystems," Am. Nature, Vol. 124 (1984), pp 293-298.

49. Lane, P. and R. Levins, "The Dynamics of Aquatic Systems. 2. The Effects of Nutrient Enrichment on Model Plankton Communities," Limnol. Oceanogr., Vol 22, No. 3 (1977), pp 454-471.

50. Cramer, G. M., R. A. Ford, and R. L. Hall, "Estimation of Toxic Hazard--A Decision Tree Approach," Food Cosmet. Toxicol., Vol 16 (1978). pp 255-276.

51. Burman, N. P. and J. S. Colbourne, "Techniques for the Assessment of Growth of Micro-Organisms on Plumbing Materials Used in Contact With Potable Water SupplieS," J. Appl. Bacteriol., Vol 43 (1977), pp 137-144.

52. Colbourne, J. S. and D. A. Brown, "Dissolved Oxygen Utilization as an Indicator of Total Microbial Activity on Non-Metallic Materials in Contact With Potable Water," J. Appl. Bacteriol., Vol 47 (1979), pp 223-231.

53. Kenaga, E. E., "Predictability of Chronic Toxicity From Acute Toxicity of Chemicals in Fish and Aquatic Invertebrates," Environ. Toxicol. Chem., Vol 1 (1982), pp 347-358.

54. Janardan, S. K., C. S. Olson and D. J. Schaeffer, "Quantitative Comparisons of Acute Toxicity of Organic Chemicals to Rats and Fish," Ecotox. Environ. Safety, Vol 8 (1984), pp 531-539.

55. Cassarett, L. J. and J. Doull, Toxicology: The Basic Science of Poisons (Macmillan, 1975).

56. Tatken, R. L. and R. J. Lewis, Sr., eds., Registry of Toxic Effects of Chemical Substances, 1981-82 edition (National Institute for Occupational Safety and Health, U.S. Department of Health and Human Services, Cincinnati, OH, June 1983).

57. Office of Science and Technology Policy, "Chemical Carcinogens: Review of the Science and Its Associated Principles," Fed. Regist., Vol 49, No. 100 (22 May 1984), p 21646.

58. Johnston, J. and P. K. Hopke, "Estimation of the Weight-Dependent Probability of Detecting a Mutagen With the Ames Assay," Environ. Mutagen., Vol 2 (1980), pp 419-424.

59. Ross, R. H., H. T. Kemp, M. G. Ryon, A. S. Hammons and J. T. Ensminger, Chemicals Tested for Phytotoxicity, ORNL/EIS-115, Vols 1 and 2 (Oak Ridge National Laboratory, 1979).

60. Kimball, K. D. and S. A. Levin, "Limitations of Laboratory Bioassays: The Need for Ecosystem-Level Testing," Bioscience, Vol 35, No. 3 (1985), pp 165-171.

61. Heinze, J. E. and N. K. Poulsen, "The Optimal Design of Batteries of Short-Term Tests for Detecting Carcinogens," Mutat. Res., Vol 117 (1983), pp 259-269.

62. Schaeffer, D. J., E. W. Novak, E. E. Herricks and K. G. Janardan, "Correlation Between Tiers in a Test Battery: Theory and Application to Mixed Aerosol Effects in an Ecological Risk Assessment," paper presented at the International Symposium on Environmental Risk Assessment of Chemicals--Fate and Effects of Chemicals in the Environment, Tokyo, Japan (18-30 May 1985).

63. Cooper, J. A., R. Saracci and P. Cole, "Describing the Validity of Carcinogen Screening Tests," Brit. J. Cancer, Vol 39 (1979), pp 87-89.

64. Purchase, I. F. H., "An appraisal of Predictive Tests for Carcinogenicity," Mutat. Res., Vol. 99 (1982), pp 53-71.

65. Bishop, Y. M. M., S. E. Fienberg and P.W. Holland, Discrete Multivariate Analysis: Theory and Practice (MIT Press, 1975).

Blank

APPENDIX A

STATISTICAL RELATIONSHIPS BETWEEN TESTS AND BATTERIES

A recent study[15] has found that, with the exception of lethality (LD50, etc.), there is virtually no toxicological data for the 60,000 compounds in commercial use or the 5 million distinct chemical entities identified in Chemical Abstracts. Given that the costs of definitive testing for one or a few toxicological endpoints can exceed $500,000 for a single compound, such data will continue to be sparse. An alternative to definitive testing utilizes the results from a battery of short-term predictive tests, and characteristics of an optimal battery have been described[13,61,62]. Two fundamental assumptions in the use of such a battery are (1) dependence between each predictive test and the definitive test and (2) independence between predictive tests. A quantitative measure of dependence between predictive tests would make it possible to develop optimal batteries of tests even for dependent assays.

We are concerned with the screening of a large number of chemicals, N, of which N1 are found to be toxic under definitive testing (Test 1 in Table A-1) and N2 are found to be nontoxic. Of the N1 true toxic compounds, 'a' are found to be toxic by a battery of screening tests (Test 2 in Table A-1) and 'b' are found to be nontoxic by these tests. Of the N2 nontoxic compounds, 'c' are reported to be toxic, and 'd' are reported nontoxic by the screening battery. This arrangement of outcomes from two batteries is given in Table A-1. Define the outcomes 'T' as "definitive" and the outcomes 't' as "predictive." For this arrangement, "prevalence" is (a + b)/N, "sensitivity" is a/(a + b), and "specificity" is d/(c + d)[63].

TABLE A-1
Outcomes From Two Assays (Tests or Batteries)[64]

		Test 2		
		t+	t-	Total
Test 1	T+	a	b	N1
	T-	c	d	N2
	total	a + c	b + d	N

Let A denote the property that the compound is positive by the first assay and B denote the property that the compound is positive by the second assay. Then, if the presence of property A promotes that of property B, we say that there is a positive correlation between the properties A and B. If the presence of property A impedes the property B, we say that the correlation between the two properties is negative. Finally, the properties are independent if the presence of the one does not influence that of the other[13,62].

Let P_1 be the probability that a given compound has property A (e.g., neurological toxicity); P_2 that the compound has property B (e.g., mutagenicity); P_{12} that the compound has both the properties A and B. The properties are independent if $P_{12} = P_1 P_2$. There is positive correlation between the properties A and B if $P_{12} > P_1 P_2$ and there is a negative correlation if $P_{12} < P_1 P_2$.[13]

If the events a, b, and ab are not independent, we can introduce instead of them three-wise independent events with probabilities such that:

X_i = 1 if the compound is positive for the ith assay, else $X_i = 0$

Then x_i for i = 1, 2 are correlated Bernoulli random variables with:

P(compound is positive for the ith assay)
$= P(X_i = 1) = p_i$ for i =1, 2

P(compound is positive for both assays)
$= P(X_1 = 1, X_2 = 1) = P_{12}$.

Estimates of the parameters p_1, p_2 and p_{12}, based on a random sample of size N, and obtained by the method of maximum likelihood, are given by:

$$p_1 = \sum X_1/n = (a + b)/N,$$

$$p_2 = \sum X_2/n = (a + c)/N \qquad \text{(A1)}$$

$$p_{12} = \sum X_1X_2/n = a/N$$

where the sum is taken over the range 1 to N.

From these definitions, a formal measure of the true correlation between the properties A and B is given by[13]:

$$\text{Corr}\ (A,B) = \text{Corr}\ (X_1, X_2) = (p_{12} - p_1p_2)/[p_1p_2(1 - p_1)(1-p_2)]^{1/2} \qquad \text{(A2)}$$

Like the familiar product-moment correlation, r, values of Corr (A,B) lie between -1 (when $p_{12} = 0$, then $p_{12} = p_1$) and +1 (when $P_1 = P_2$, then $P_{12} = P_1$). Significance levels for the correlation between batteries are obtained from the distribution of χ^2 since $\text{Corr}^2(A,B) = \chi^2/N$ with 1 degree of freedom (df).

This derivation of the correlation between batteries provides a general relationship for the correlation among K predictive tests. If Test 1 and Test

2 in Table A-1 are outcomes from two predictive tests, Corr(A,B) is a formal measure of test independence.

$$Corr^2(A,B) = (ad-bc)^2/[(a+b)(a+c)(c+d)+(b+d)] \equiv \chi^2/N \qquad (A3)$$

The correlation, $Corr^2(A,B)$ with 1 df, is thus formally related to the χ^2 criterion for independence between a pair of predictive tests.[64] However, the data from K > 2 tests conforms to multinomial sampling and affords a K x K table. Since it is rare to have enough data to calculate χ^2 for the K x K table, Heinze and Poulsen[61] calculated χ^2 values for all possible pairs of tests. They assumed that independence of all pairs represents independence of all tests. This key assumption is incorrect. While independence in the expanded K x K table implies independence in the condensed (2 x 2) tables, the converse does not hold since the latter can be derived from a larger table with more complex structure[65].

Defining L as the number of positive responses required to classify a compound as an environmental toxicant, the correlation among K > 2 predictive tests in a battery is also given by (A3), but with K-(L-1) df. The differences in estimating the correlation between the two batteries and among the K > 2 tests in a battery are: (1) cell frequencies for batteries are obtained as the union of the K tests while cell frequencies for K tests are obtained as the intersection of the L positive tests; and (2) there is 1 df associated with the correlation between batteries and K-(L-1) df associated with the correlation of tests within a battery. These relationships are clarified in the numerical examples that follow.

EXAMPLES

This example[62] considers the induction of micronuclei (MCN) in Tradescantia by chemicals. We want to know whether mutagenicity in Salmonella predicts induction of MCN in Tradescantia. Omitting results from overdosed experiments, the results for 30 chemicals were a = 12, b = 5, c = 4, d = 9. The proportion of true positives is 12/17 and the proportion of true negatives is 9/13. The correlation between assays is:

$$\text{Corr}(\underline{\text{Salmonella}}, \text{MCN}) = (ad-bc)/[(a+b)(a+c)(b+d)(c+d)]^{1/2}$$
$$= [12(9) - 5(4)]/[(12+5)(12+4)(5+9)(4+9)]^{1/2}$$
$$\chi^2 = N\,\text{Corr}^2 = 4.69 \quad (p < 0.5).$$

Results from 26 compounds tested in Tradescantia were compared with mammalian teratogenicity data, giving: a = 8, b = 5, c = 6, d = 7. The proportion of true positives is 8/13 and the proportion of true negatives is 7/13. The correlation is not significant:

$$\text{Corr}(\text{Teratogenicity, MCN}) = (56-30)/[13(14)(13)12]^{1/2} = 0.154.$$

Assume that we have Salmonella and teratogenicity data (predictive tier) for 30 compounds, and that 15 of these are ultimately positive in the MCN test (definitive tier). Neither test in the predictive tier, alone, is good at predicting MCN. If one positive test is enough to classify a compound as positive, the correlation between tiers is not significant:

$a = 15[12/17 + 8/13 - (12/17)(8/13)] = 13$; $b = 15 - 13 = 2$;
$d = 15[(9/13)(7/13)] = 6$; $c = 15 - 6 = 9$.

$$\text{Corr}(\text{Predictive, MCN}) = [13(6)-2(9)]/[(13+2)(13+9)(2+13)(2+6)]^{1/2} = 0.30.$$

However, if two positives are required, then a = 7, b = 8, c = 2, d = 13, and the correlation between tiers, 0.36, is significant.

Consider one predictive assay that has a sensitivity = specificity = 0.9. If the true proportion of toxicants is 0.1, and N = 1000, then a = 90, b = 10, c = 90, d = 810. Using (A3),

$$\text{Corr}^2(A,B) = [90(810)-10(90)]^2/[100(180)(900)(820)] = 0.390,$$

and the correlation between the predictive and definitive tests, 0.62, is significant. A second predictive assay, also with sensitivity = specificity = 0.9, is added to the predictive battery. Between them, the predictive tests classify compounds as follows: a = 81, b = 19, c = 10, d = 890. In general, if S_1 and S_2 are the sensitivity and specificity of the battery at the point of L positives, cell entries are given by: $a' = \text{integer}\,(N1\ S_1^L + 0.5)$, $b' = N1 - a'$, $d' = \text{integer}\,(N2\ S_2^L + .5)$ and $c' = Ns - d'$.

When these two tests comprise the predictive battery, either one positive response or two positive responses can be required to classify the compound as an environmental toxicant. If one positive response is enough, the theoretical sensitivity of the battery is 0.99 and the specificity is 0.81, so that a = 99, b = 1, c = 171, d = 729[61]. Using (A3), the correlation between batteries is 0.54, but with 2 df. Both correlations are significant. If two

positive results are required, the correlation between batteries is 0.84. The correlation between predictive assays, 0.69 (1 df), is obtained from (A3) using $a' = \text{integer}\,[100[0.81^2) + 0.5] = 66$, $b' = 34$, $d' = [900(0.99^2) + 0.51] = 882$, $c' = 18$.

Tables A-2 and A-3 give the correlations for batteries of various size for assays with sensitivity = specificity = 0.65. Comparing these correlations with those given above shows that the magnitude and sign of a correlation within and among batteries depends on the sensitivity and specificity of each assay. For example, the correlation between three assays with sensitivity = specificity = 0.9 is positive, but is negative when sensitivity = specificity = 0.65; both correlations are significant. Comparing these results with those for a two-test battery shows that the magnitude of a correlation also depends on the number of assays in a battery and the number of positive responses used as a criterion.

The results in Tables A-2 and A-3 clearly show that batteries can be correlated whether or not the individual assays are independent. Comparing the foregoing examples and Table A-4 with Tables A-2 and A-3, it is seen that the magnitude of a correlation for a given battery size is smaller for assays having relatively low, as compared with relatively high, sensitivity and specificity. Further, the correlation relationship is nonlinear; it rises to a maximum and then falls off nonsymmetrically (Table A-3). The optimal decision criterion for this battery, based on maximizing both the sensitivity and specificity, is eight positive responses, whereas the maximum correlation requires nine positive results. The table clearly shows that certain decision criteria (e.g., 14 positives) imply poor predictability since the batteries are uncorrelated.

The true proportion of toxicants in the sample affects the magnitude, hence significance, of correlations between both batteries and tests. This is shown in Table A-4 for a battery comprised of five tests with nominal sensitivity = specificity = 0.9, for true proportions of toxicants of 0.1, 0.3, 0.5, 0.7, 0.9. Maximum correlation occurs with three positive tests for all proportions of toxicants, in agreement with Heinze and Poulsen[61]. This contrasts with the results in Table A-2 for sensitivity = specificity = 0.65, where our optimal decision rule is four positive tests. Therefore, in addition to the effect of test battery size[61], our optimal decision rule is also a function of sensitivity and specificity, but not sample size.

Table A-2
Correlation for a Battery of Short-Term Tests[13] Each of Which has Sensitivity = Specificity = 0.65 and True Proportion of Toxicants = 0.1

Number		Sensit-	Specif-					Correlation	
Tests	Positive	vity	icity	a	b	c	d	Tiers	Tests
1	1	0.65	0.65	65	35	315	585	0.185	
2	1	0.878	0.422	88	12	520	380	0.185	0.185
	2	0.422	0.878	42	58	110	790	0.248	-0.035
3	1	0.957	0.274	96	4	653	247	0.162	0.162
	2*,**	0.718	0.718	72	28	254	646	0.280	0.021
	3	0.274	0.957	27	73	39	861	0.273	-0.098
4	1	0.985	0.178	99	1	740	160	0.136	0.136
	2	0.874	0.563	87	13	393	507	0.260	0.050
	3**	0.563	0.874	56	44	113	787	0.347	-0.098
	4	0.178	0.985	18	82	13	887	0.286	-0.079
5	1	0.995	0.116	100	0	796	104	0.113	0.113
	2	0.946	0.428	95	5	515	385	0.232	0.058
	3*	0.765	0.765	77	23	211	689	0.354	-0.062
	4**	0.428	0.946	43	57	49	851	0.389	-0.131
	5	0.116	0.995	12	88	4	896	0.276	-0.050
7	4*	0.8	0.8	80	20	180	720	0.410	-0.109
	5**	0.532	0.944	53	47	50	850	0.468	-0.150
9	5*	0.828	0.828	83	17	155	745	0.463	-0.135
	6**	0.609	0.946	61	39	48	852	0.536	-0.160
11	6*	0.851	0.851	85	15	134	766	0.508	-0.147
	7**	0.668	0.950	67	33	45	855	0.590	-0.162
13	7*	0.871	0.871	87	13	116	784	0.552	-0.147
	8**	0.716	0.954	72	28	42	858	0.636	-0.162
15	8*	0.887	0.887	89	11	102	798	0.592	-0.145
	9**	0.750	0.958	75	25	38	862	0.671	-0.158

*Optimal decision criterion given by Heinze and Poulsen[61].

**Optimal decision criterion based on maximum correlation between batteries[13,62].

Table A-3
Relationship of Correlation[13] to Decision Criterion for 15 Assays, Each of Which has Sensitivity = Specificity = 0.65 and True Proportion of Toxicants = 0.1

Criterion	Sensitivity	Specificity	Correlation		χ^2
			Tiers	Tests	Tests*
1	1.0000	0.0016	0.011	0.011	0.11
2	0.9999	0.0141	0.038	0.	0.
3	0.9999	0.0617	0.081	0.	0.
4	0.9995	0.1726	0.147	0.011	0.11
5	0.9971	0.3519	0.227	-0.017	0.30
6	0.9875	0.5642	0.332	-0.061	3.70
7	0.9578	0.7548	0.460	-0.100	10.05
8	0.8867	0.8867	0.593	-0.145	21.10
9	0.7548	0.9578	0.671**	-0.158	25.06
10	0.5642	0.9875	0.657	-0.115	13.17
11	0.3519	0.9971	0.544	-0.056	3.20
12	0.1726	0.9995	0.395	-0.24	0.56
13	0.0617	0.9999	0.233	-0.010	0.11
14	0.0141	0.9999	0.095	0.	0.
15	0.0016	1.0000	0.000	0.	0.

*Values of $\chi^2 > 3.84$ are significant.
**Maximum correlation.

Table A-4
Dependence of Correlation on True Proportion (P)
of Toxicants for Five Tests, N = 1000

No. Pos. Tests	Sensitivity/ Specificity (P)	a	b	c	d	Correlation Tiers	Correlation Tests	χ^2 (Tests)
1 (0.1)	0.999/0.59	100	0	369	531	0.354	0.354	125.8
2	0.999/0.919	100	0	73	827	0.728	0.593	351.8
3	0.991/0.991	99	1	8	892	0.952	0.867	753.0
4	0.919/0.999	92	8	1	899	0.949	0.803	645.8
5	0.590/0.999	59	41	1	899	0.743	0.188	35.6
1 (0.3)	0.999/0.59	300	0	287	413	0.549	0.549	301.5
2	0.999/0.919	300	0	57	643	0.878	0.784	614.8
3	0.991/0.991	297	3	6	694	0.978	0.936	877.3
4	0.919/0/999	276	24	1	699	0.940	0.788	621.3
5	0.590/0.999	177	123	1	699	0.705	0.196	38.7
1 (0.5)	0.999/0.59	499	1	205	295	0.644	0.644	414.7
2	0.999/0.919	499	1	40	460	0.920	0.852	726.1
3	0.991/0.991	496	4	4	496	0.984	0.984	898.7
4	0.919/0.999	460	40	1	499	0.920	0.740	547.6
5	0.590/0.999	295	205	1	499	0.644	0.177	31.6
1 (0.7)	0.999/0.59	699	1	123	177	0.705	0.705	497.1
2	0.999/0.919	699	1	24	276	0.940	0.886	785.5
3	0.991/0.991	694	6	3	297	0.978	0.936	877.3
4	0.919/0.999	643	57	0	300	0.878	0.650	422.8
5	0.590/0.999	413	287	0	300	0.549	0.141	20.1
1 (0.9)	0.999/0.59	899	1	41	59	0.743	0.743	553.3
2	0.999/0.919	899	1	8	92	0.949	0.896	803.6
3	0.991/0.991	892	8	1	99	0.952	0.867	753.0
4	0.919/0.999	827	73	0	100	0.728	0.446	199.2
5	0.590/0.999	531	369	0	100	0.354	0.087	7.5

APPENDIX B

ESTIMATOR FOR RAT ORAL LD50*

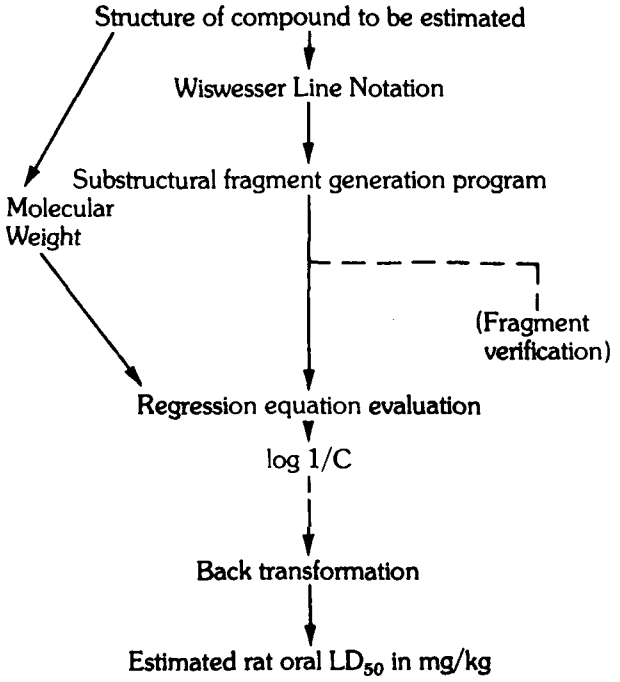
*Used with permission of Mr. K. Enslein, President, Health Designs, Inc.

Estimator for Rat Oral LD_{50}

Description

Rat Oral LD_{50} in mg/kg is estimated by evaluation of a regression equation. The estimating equation uses substructural fragments and molecular weight as independent parameters.

The computer-based estimation process is shown below:



Data used in development of estimating equation

The chemicals were drawn from the April 1979 tape version of the Registry of Toxic Effects of Chemical Substances (RTECS), published by the National Institute of Occupational Safety and Health (NIOSH). This compendium lists the most toxic effects of those reported in the literature. The tape is reasonably free of errors, and represents the most comprehensive compilation of LD_{50} data presently available.

Number of chemicals used in development of estimating equation

The April 1979 version of RTECS included approximately 3,600 compounds with rat oral LD_{50} values. A randomly selected subset of 2,000 of these formed the basis of the structure-activity equation. Removal of duplicates, some outliers, and restriction of range reduced the actual number of compounds to 1,851.

Estimation range

The toxicity of compounds used for the model development had a range of log 1/C of 1.25 to 4.75, where $1/C = (1{,}000 \times MW)/LD_{50}$ expressed in mg/kg. The molar form is preferred and used in all computations.

Terms in estimating equation

93	substructural fragments
1	log molecular weight
1	constant
95	

Performance characteristics

567 compounds were randomly selected from the population of 3,600. These compounds were *not* used in the development of the estimating equation. Log 1/C values were then calculated for these compounds by evaluating their rat oral LD_{50} with the estimating equation. These predicted values were then subtracted from the corresponding values in RTECS, and residuals obtained. Quantiles and simple statistics were then obtained for various ranges of predicted log 1/C. The resulting tables are shown below:

Predicted log 1/C	N	Quantiles							
		1	5	10	25	75	90	95	99
1.0-1.5	3	.12	.12	.12	.12	1.74	1.74	1.74	1.74
1.5-2.0	78	1.37	1.05	.71	.42	.26	.54	.79	1.00
2.0-2.5	224	1.34	.87	.67	.45	.34	.76	1.03	1.74
2.5-3.0	143	1.99	1.25	.78	.45	.36	.86	1.45	2.71
3.0-3.5	97	1.80	1.20	.81	.46	.52	1.17	1.56	2.51
3.5-4.0	18	1.63	1.63	1.60	1.07	.50	1.48	1.64	1.64
4.0-4.5	4	1.02	1.02	1.02	1.02	.85	.85	.85	.85

	Simple Statistics				
Predicted log 1/C	Avg.	Med.	S.D.	Min.	Max.
1.0-1.5	.72	.55	.94	−.12	1.74
1.5-2.0	.059	.12	.66	−1.37	1.00
2.0-2.5	−.015	−.059	.58	1.49	1.94
2.5-3.0	.017	.07	.77	2.33	2.76
3.0-3.5	.081	.067	.81	−1.80	2.51
3.5-4.0	.21	.35	.98	−1.63	1.64
4.0-4.5	.082	.077	1.07	−1.02	.85

To show how we estimate the confidence bounds for a predicted value, an example will be helpful. Suppose the predicted log 1/C value is 2.7. Then entering the table at the line "2.5 - 3.0", the 10th and 90th quantiles have the values −.78 and +.86, respectively. Thus 10% of the residuals have values more negative than −.78 and 10% have values more positive than .86. Maximum deviations for other quantiles can be determined in a similar fashion.

The estimating equation

The first 30 most important terms of the equation are shown in Figure 2. The "F" value is a variance ratio which indicates the relative importance of each of the terms of the equation.

Regression Statistics

log 1/C Mean	2.533
Standard Deviation	0.755
Range	1.25 - 4.73
R^2 (multiple correlation coefficient)	0.449
Residual mean square	0.33
Standard error of estimate	0.57

Figure 2

LD_{50} MODEL

TERMS LISTED IN ORDER OF DECREASING IMPORTANCE

WLN Key #	Frequency In 1851 Compounds	Description	Regression Coefficient	F
11	114	More Than 1 Sulphur Atom	.821	150.7
293	22	Sn	1.175	78.6
5	158	Terminal Oxygen (Not Carbonyl)	.458	62.5
		Log Molecular Weight	.681	53.5
10	185	1 Sulphur Atom	.362	52.2
309	61	Substituent Carbamate	.548	38.5
20	223	Alkyl Chain $(CH_2)n$ or $CH_3(CH_2)n-1$, where n 3-9	−.250	31.6
341	79	Aromatic Nitro	.399	31.3
256	9	Hg	1.114	29.3
28	153	One NH Group Chain Fragment	.334	29.1
315	161	Haloalkane	.323	21.2
104	233	1 Heteroatom in 1 Ring	.202	18.6
166	106	Chain Dialkylamino	.255	16.4
250	2	Pb	1.164	16.0
111	26	More Than 1 Single Heterocyclic Ring	.574	15.6
107	56	1 Heteroatom In More Than 1 Ring	.477	15.4
30	102	One NH_2 Group Chain Fragment	.236	14.5
36	205	More Than One O Group Chain Fragment	.211	14.4
130	99	Bilinkage	.271	13.9
50	280	Substituent Generic Halogen	.212	12.6
37	195	One OH Group Chain Fragment	−.163	12.2
193	14	Substituent Sulfonamide	−.561	12.0
34	54	Unusal Carbon Atom Chain Fragment	.278	11.5
26	30	Fluoride	.435	11.4
322	4	Aziridine	.934	10.1
330	16	Fused Aromatic-Unsaturated Lactone	.539	10.1
344	83	Geminal-Dihaloalkane	−.315	10.0
201	4	Substituent N-Nitro	−.934	9.0
112	65	1 Single Carbocyclic Ring	.321	9.0
282	13	Si	.484	8.6
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Constant			.552	